BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

FIELDS

VIROLOGY

Third Edition

Volume 2

Editors-in-Chief

Bernard N. Fields, M.D.
Chairman, Department of Microbiology and Molecular Genetics
Harvard Medical School
Boston, Massachusetts

David M. Knipe, Ph.D.
Department of Microbiology and Molecular Genetics
Harvard Medical School
Boston, Massachusetts

Peter M. Howley, M.D.
Chairman, Department of Pathology
Harvard Medical School
Boston, Massachusetts

Associate Editors

Robert M. Chanock, M.D.
Laboratory of Infectious Diseases
National Institute of Allergy and Infectious Disease
National Institutes of Health
Bethesda, Maryland

Joseph L. Melnick, Ph.D., Sc.D.
Department of Virology and Epidemiology
Baylor College of Medicine
Texas Medical Center
Houston, Texas

Thomas P. Monath, M.D.
Ora Vax, Inc.
Cambridge, Massachusetts

Bernard Roizman, Sc.D.
Department of Molecular Genetics and Cell Biology
University of Chicago
Chicago, Illinois

Stephen E. Straus, M.D.
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Bethesda, Maryland

Lippincott - Raven
PUBLISHERS
Philadelphia • New York

Lippincott-Raven Publishers, Philadelphia, 227 East Washington Square, Philadelphia, PA 19106

Made in the United States of America

Library of Congress Cataloging-in-Publication Data

Fields Virology / editors-in-chief, Bernard N. Fields, David M. Knipe, Peter M. Howley ; associate editors, Robert M. Chanock . . . [et al.].—3rd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-7817-0253-4 (alk. paper)
1. Virology. I. Fields, Bernard N. II. Knipe, David M. (David Mahan), III. Howley, Peter M.
[DNLM: 1. Viruses. 2. Virus Diseases. QW 160 F463 1995]
QR360.V5125 1995
616'.0194—dc20
DNLM/DLC
For Library of Congress 95-12782

9 8 7 6 5 4 3 2 1

same place in the viral sequence. Analysis of the effects of mutations near the ends of the LTRs on integration *in vivo* and *in vitro* show that this specificity is provided both by recognition of a specific sequence containing the inverted repeat at each end and by proximity of this sequence to the ends of the viral DNA (141,774). This sequence, sometimes called *att,* probably forms the only signal required in *cis* by the viral integration machinery. Mutational analysis has narrowed the sequence required to as few as 6 bases from the ends of the LTR (92,527,660).

From the perspective of the cellular target, most approaches to the problem suggest that integration is a much more random process. Comparison of cellular sequences flanking a modest number of independent integration events (12 to 20 in most cases studied) reveal no common sequence which might serve as a cellular target. Similarly, analysis of the distribution of integration events into target DNA *in vitro* is not suggestive of any specificity for a particular site (81), although it may be possible to define loose "consensus" sequences by analyzing large numbers of integration joints (223). More incisive analyses using PCR to determine the pattern of integration into small defined regions of target DNA reveals that integration targets can be found in all regions examined, but that there is a decidedly nonrandom pattern of usage of specific sites within any region (395,597). This pattern differs from one virus to another (Y. Kitamura and J. M. Coffin, unpublished observations) and seems to reflect the interaction of the integrase system with local structural features. Incorporation of a DNA target into chromatin does not block integration but does alter its specificity in a striking way, such that integration targets appear with 10-base periodicity where nucleosomes are present (599). Similarly, integration is not inhibited by C-methylation of DNA, and this modification can even create strong target sites (395). Strong target sites for integration *in vitro* can also be created by introducing bends into the target DNA (596). The role of such features in the infected cell remains to be examined.

Another approach to the issue of integration specificity is to compare sites of integration selected for insertion into the same general region, e.g., in tumors induced by activation of proto-oncogenes. This type of analysis can be complicated by selection for integration in specific regions by effects on cell growth. In B-cell lymphomas induced by ALV inoculation of chickens (see below), the large majority of tumors have an ALV provirus inserted within the first intron of the c-*myc* protooncogene (259,628). Since proviruses have not been found in this region in any other experiment, there is no reason to believe that integration into it is especially favored. Their presence is a consequence of selection of cells transformed by provirus alteration of c-*myc* expression. Selection for proviruses integrated in specific regions can also be more subtle, as a consequence of reduced growth factor dependence of tissue culture cells (746), for example. For this reason, inferences regarding integration targeting are reliably drawn only from analysis of newly infected cells.

Several reports that suggest a tendency for integration to occur in regions of DNA that might tend to be transcriptionally active, e.g., regions characterized by relatively "open" chromatin structure. Proviruses cloned from MLV-infected cells were often found to be integrated in the vicinity of active genes or DNase-sensitive sites which often mark actively transcribed genes (631,648,772), and the *myc*-associated ALV-proviruses mentioned above showed a tendency to be located in proximity to one of five DNase-hypersensitive sites in c-*myc* (628). Another level of specificity has been reported for a fraction of integrations of ALSV into avian cell DNA (662), which suggested that some 20% of integrations are into one of about 1,000 specific sites.

A more recent analysis which makes use of PCR to detect and localize single integration events in large cultures of cells yields a rather different picture of the specificity issue (807). Targets for ALSV integration in newly infected cells were found to be distributed very much like targets *in vitro*. All tested regions of the genome contained integration sites, with very local hot spots. On average, the frequency of targets per region was like that expected on a purely random basis, suggesting the absence of any strong regional targeting.

Mechanism of Integration

Studies of the mechanism of integration have been made possible by the availability of powerful *in vitro* systems. The first of these consisted simply of extracts of cells made at a time (about 12 to 24 hours after infection) when viral DNA synthesis is complete but integration is still ongoing. If an appropriate target DNA is added to such an extract (either from nuclei or cytoplasm), then integration of the viral DNA mediated by preintegration complexes in the extract can be readily detected. Suitable detection methods include the use of a selectable marker and cloning after integration into a suitable vector (81,224), Southern blotting to reveal integration into a closed circular DNA target (206,240,426), and PCR using primers derived from the target DNA and the LTR (395,599). These reactions are quite efficient. In all cases, detailed analysis of the reaction products shows that the *in vitro* reaction is correct. There is a loss of the appropriate number of bases from the end of the viral DNA and a duplication of the correct number of bases of cellular sequence flanking the provirus.

Success with reactions from preintegration complexes led to the development of model reactions using IN protein purified from virions or produced by recombinant techniques and simple labeled double-stranded oligonucleotides as substrates (94,373,378). Although the purified systems are much less efficient than reactions using preintegration complexes, and their biochemical properties (such as divalent cation requirement) are somewhat different, they have permitted identification of IN as the source of all nec-

essary catalytic activity, as well as in-depth analysis and understanding of the integration mechanism. In these reactions, incubation of suitable substrate oligonucleotides with IN, followed by analysis by electrophoresis in denaturing polyacrylamide gels, reveals two major classes of product: a molecule two bases shorter than the substrate, representing the product of the cleavage reaction which removes the 3′ dinucleotide and a heterogeneous collection of larger molecules resulting from the strand transfer reaction in which the 3′ end of the substrate oligonucleotide is integrated into an internal position on another (target) molecule. In the simplest system, substrate and target are the same; however, integration into other target molecules such as plasmid DNA can also be assayed. This provides the basis for some rapid, quantitative assays for integrase activity (149,524).

Using these approaches, the following information has been gleaned regarding the integration process.

First, concordant with the rationale, integration in extracts of infected cells is mediated by the preintegration complexes; the activity of complexes containing only linear DNA molecules implies the linear form of DNA as the important intermediate, a conclusion well substantiated by subsequent biochemical analysis. At least in the case of MLV and HIV, structures purified from soluble cellular components by sedimentation or chromatography are still active in integration, implying that no soluble cell factors are required (65,206). Indeed, active HIV complexes containing only DNA and IN can be isolated (208), indicating that other viral proteins, though often present, may also be dispensable.

Second, the reaction carried out by the MLV complexes is limited to integration into the added target. No side reactions, such as circle formation or autointegration, are seen (82,240). In the case of ALSV and HIV, however, the situation seems to be somewhat more complex. Circular products due to integration of the viral DNA into itself as well as single LTR circles are found among the reaction products in high yield (Fig. 12) (207,426,427), although their formation can be suppressed by appropriate reaction conditions (427). Presumably, there are specific mechanisms to block autointegration *in vivo*, although small amounts of one- and two-LTR circles, as well as the circular products arising from integration of viral DNA molecules into themselves, can be found in infected cells (666). Recent evidence suggests that a host cell factor, tightly bound to the MLV complex, may suppress autointegration in this virus (R. Craigie, personal communication). How this important aim is accomplished in other viruses remains unclear. In the case of ALSV, the DNA in cytoplasmic preintegration complexes is incomplete (427). Perhaps there is a block to some event required for completion of DNA synthesis (such as the plus-strand jump) that causes a delay until the complex is in the nucleus where integration targets are abundant.

Third, most of the 3′ ends of the viral DNA in the plexes have already been cleaved by IN to remove tl minal two bases, while the 5′ end of each strand i cisely at the site of its initiation, and therefore ha modified only by removal of the primer. The ends thu the sequence

5′ AATG.....CA 3′

3′ AC.....GTAA 5′

Since the majority of unintegrated forms can ha structure some time prior to the integration reacti formation must not be tightly coupled kinetically rest of the integration reaction. However, cleavages ends of a molecule are coupled to one another; mu in MLV at one *att* site block cleavage at both (529

The preintegration complex reaction joins the pre formed 3′ end to a 5′ end of the target DNA (82,24(5′ ends of the viral DNA are unaltered and remain un This result strongly implicated a linear molecule as tegration intermediate rather than a two-LTR circu termediate, as was previously thought. This conclusi confirmed by the suitability of small oligonucleoti substrates for integrase.

Fourth, mutational analysis implies that the cleava strand transfer reactions take place at the same acti (195,372,414,759), requiring the highly conserved I E motif. While the coincidence of the two differen tions at the same site might at first seem contradicto underlying mechanism of the two reactions is, in fa same (196,759) (Fig. 16). In both cases, the reacti volves a direct attack on a phosphate group by a hy group, leading to exchange of an internucleotide b the usual cleavage reaction, the OH donor is water ever, other molecules, such as glycerol or the 3′ OH same DNA (leading to a cyclic dinucleotide produ also participate (196). In the strand transfer reacti OH donor is the 3′ end of the newly cleaved DNA ing to a direct transesterification. Thus, the integrat action is concerted, rather than involving separate age and ligation reactions.

Fifth, there is no obvious requirement for a par sequence or structure of the target DNA. Indeed, a above, both incorporation of DNA into chromatir and extensive methylation of dC residues (395)—n cations associated with reduced transcriptional acti can create strong target sequences for integration *i* Their *in vivo* effects remain to be tested.

Finally, the reaction proceeds efficiently in the a of ATP or any other ad... energy-generating (81,94). This independence is consistent with the c breakage-rejoining reaction. Also consistent is t servation that IN can catalyze the reverse reacti ferred to as *disintegration*. If provided with a DN ecule resembling the integration intermediate, I

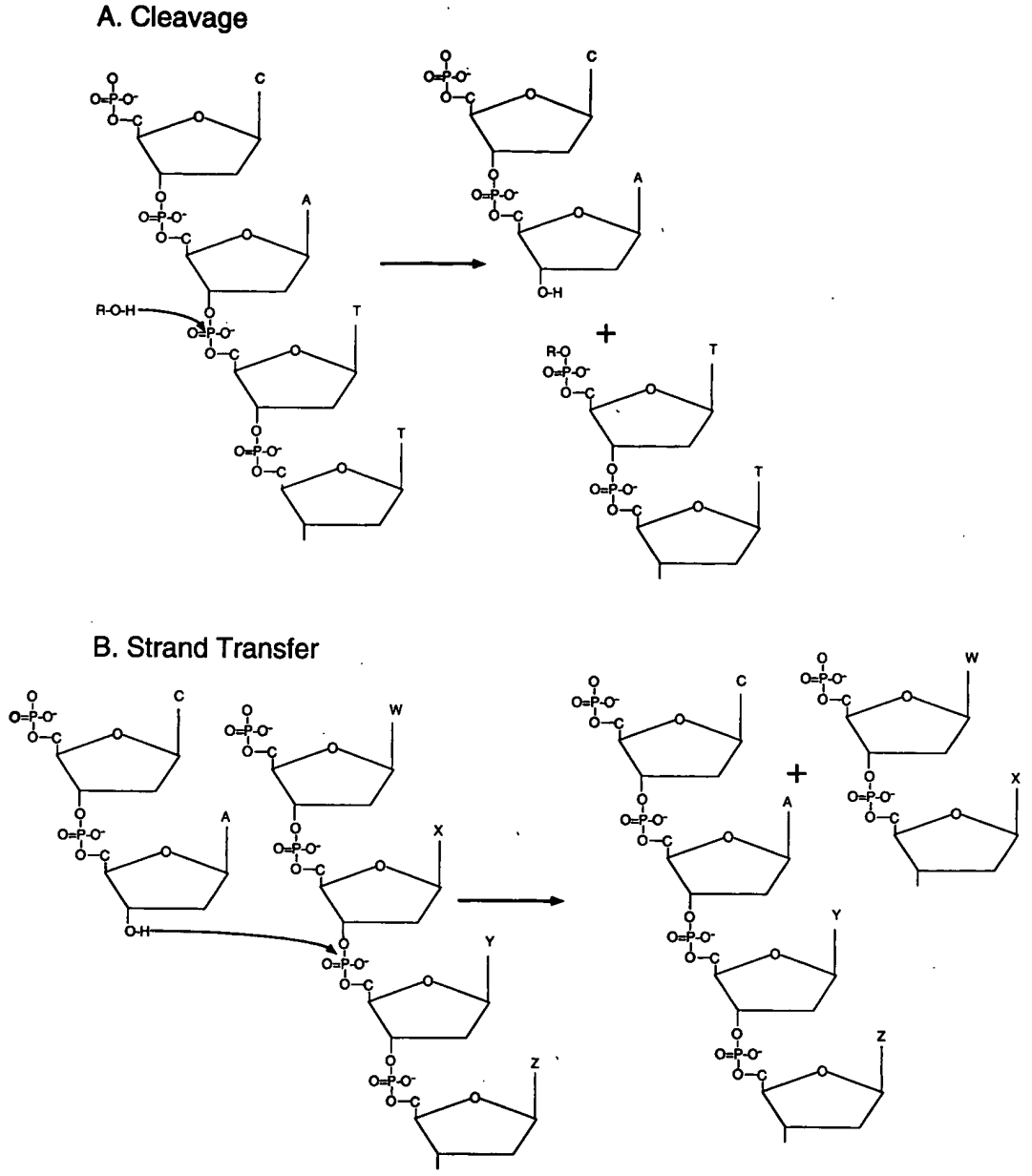

FIG. 16. IN-mediated cleavage and strand transfer. Note that the underlying mechanism (OH-mediated attack on an internucleotide phosphate bond) is the same for both reactions (196,775).

separate it into two molecules, resealing the adjacent nick in the process (120).

All of these considerations have led to formulation of the pathway shown in Fig. 17.

1. Following viral DNA synthesis, the core structure containing linear DNA, and the CA, IN, and possibly RT and NC proteins, enters the nucleus.
2. The 3′ terminal two bases at either end are removed by the cleavage reaction of IN, leaving a 3′ OH end. This reaction may occur before entry into the nucleus.
3. The strand transfer reaction simultaneously joins the two ends of the viral DNA to cellular DNA about half a turn of the helix apart, with the precise spacing determined by the geometry of the IN multimer.
4. A cellular DNA repair system fills in the resulting gap in the molecule, displacing the two mismatched bases at the 5′ end of the provirus and ligating the remaining ends. This gap repair of the initial staggered joints generates the characteristic duplication of cell DNA flanking the provirus.

JOURNAL OF VIROLOGY, Nov. 1997, p. 8429–8436
0022-538X/97/$04.00+0

Vol. 71, No. 11

Adeno-Associated Virus Vector Integration Junctions

ELIZABETH A. RUTLEDGE AND DAVID W. RUSSELL*

Department of Medicine and Markey Molecular Medicine Center, University of Washington, Seattle, Washington 98195

Received 29 April 1997/Accepted 31 July 1997

Vectors derived from adeno-associated virus (AAV) have the potential to stably transduce mammalian cells by integrating into host chromosomes. Despite active research on the use of AAV vectors for gene therapy, the structure of integrated vector proviruses has not previously been analyzed at the DNA sequence level. Studies on the integration of wild-type AAV have identified a common site-specific integration locus on human chromosome 19; however, most AAV vectors do not appear to integrate at this locus. To improve our understanding of AAV vector integration, we analyzed the DNA sequences of several integrated vector proviruses. HeLa cells were transduced with an AAV shuttle vector, and integrated proviruses containing flanking human DNA were recovered as bacterial plasmids for further analysis. We found that AAV vectors integrated as single-copy proviruses at random chromosomal locations and that the flanking HeLa DNA at integration sites was not homologous to AAV or the site-specific integration locus of wild-type AAV. Recombination junctions were scattered throughout the vector terminal repeats with no apparent site specificity. None of the integrated vectors were fully intact. Vector proviruses with nearly intact terminal repeats were excised and amplified after infection with wild-type AAV and adenovirus. Our results suggest that AAV vectors integrate by nonhomologous recombination after partial degradation of entering vector genomes. These findings have important implications for the mechanism of AAV vector integration and the use of these vectors in human gene therapy.

Adeno-associated virus (AAV) is a 4.7-kb single-stranded DNA virus that has been developed as a gene therapy vector (31). Only the terminal repeat (TR) sequences are required in *cis* for replication and packaging, allowing a complete replacement of viral coding sequences with foreign DNA in vectors. A major advantage of AAV vectors is their ability to stably transduce cells by integration into host chromosomes. Although integration is crucial for many gene therapy applications, the mechanism of AAV vector integration is poorly understood, and the structure of integrated proviruses has not been determined at the DNA sequence level.

Southern analysis of DNA from transduced cells maintained under selection for the vector transgene has usually demonstrated the presence of integrated AAV vector proviruses (19, 27, 30, 36, 42). However, there have also been reports of episomal vector molecules, especially with *rep*$^+$ vectors (30) or the absence of selectable markers (1, 4). Some reports of vector integration noted a predominance of concatemeric proviruses (30, 36), while others did not (11, 27, 33, 44). The reasons for these variable results are unclear and include differences in transduction protocols as well as possible effects of contaminating wild-type AAV functions.

Two-thirds of integrated wild-type AAV proviruses are found at a specific human chromosome 19 site, 19q13-qter (24, 25, 38). While this feature could prove useful in some gene therapy applications, AAV vectors have not been found to integrate at this same locus (33, 44). Site-specific AAV integration appears to be mediated by the viral Rep protein (16, 28), so the absence of the *rep* gene in most vectors can explain the lack of site-specific integration. The presence of random-sized junction fragments detected by Southern analysis of integrated proviruses suggests that vector integration sites are random (27, 30, 33, 44). However, it is possible that vector integration occurs at scattered locations within a common locus different from 19q13-qter, just as wild-type AAV integrants are found scattered throughout the chromosome 19 site-specific integration locus (23, 38).

To better understand the process of AAV vector integration, we produced an AAV shuttle vector that allowed us to recover integrated vector proviruses along with flanking human DNA as bacterial plasmids. Several integration junctions from independent, transduced HeLa cell clones were sequenced and compared. Our results show that vector integration occurs at random chromosomal sites and that none of the vector proviruses integrated as intact, full-length genomes.

* Corresponding author. Mailing address: Department of Medicine, Box 357720, University of Washington, Seattle, WA 98195. Phone: (206) 616-4562. Fax: (206) 616-8298. E-mail: drussell@u.washington.edu.

MATERIALS AND METHODS

Cell culture. Human HeLa (39) and 293 cells (17) were cultured in Dulbecco's modified Eagle medium with 10% heat-inactivated (56°C for 30 min) fetal bovine serum (HyClone, Logan, Utah), amphotericin (1.25 μg/ml), penicillin (100 U/ml), and streptomycin (100 μg/ml) at 37°C in a 10% CO_2 atmosphere. Titers of vector stocks were determined on HeLa cells by selecting for G418 (GIBCO-BRL, Grand Island, N.Y.) resistance as described previously (33) except that the G418 concentration was 1 mg of active compound per ml. Transduction of HeLa cells by AAV-SNori was carried out as when titering for G418 resistance except that resistant colonies were picked and expanded to 2×10^7 cells for isolation of genomic DNA. The multiplicities of infection were 0.17 vector particles per cell for clones 2 and 6; 1.7 for clones 1, 3, 4, 10, and 11; and 170 for clones 5, 7 to 9, and 12 to 14 (see Table 1).

Preparation of virus stocks. AAV-SNori vector stocks were prepared as follows. 293 cells were plated at a density of 8×10^6 cells/dish in 12 dishes (15-cm diameter). The next day, each dish was infected with 1.2×10^8 PFU of adenovirus type 5 (ATCC VR-5; American Type Culture Collection, Rockville, Md.) and 2 hours later cotransfected with 8 μg of pASNori2 and 32 μg of pAAV/Ad (36) by the calcium phosphate method (35). After 3 days, the cells and medium were harvested and combined, subjected to three cycles of freeze-thaw lysis in a dry ice-ethanol bath, clarified by centrifugation at 5,800 × *g* (5,500 rpm) in a Sorvall HS4 rotor for 30 min at 4°C, digested with micrococcal nuclease (68 U/ml; Pharmacia, Piscataway, N.J.) at 37°C for 1 h, treated with trypsin (50 ng/ml) at 37°C for 30 min, and centrifuged through 40% sucrose in phosphate-buffered saline in a Beckman SW28 rotor at 27,000 rpm for 16 h at 4°C. The pellets were resuspended in 8 ml of a 0.51-g/ml solution of CsCl and passed twice through a 22-gauge needle. The suspension was centrifuged in a Beckman SW41 rotor at 37,000 rpm for 20 h at 4°C. The region of the gradient containing AAV virions was collected, dialyzed against Dulbecco's modified Eagle medium through a 50,000-molecular-weight-cutoff membrane (Spectrum, Houston, Tex.), and concentrated by centrifugation in Centricon 100 filters (Amicon, Inc., Beverly, Mass.). Adenovirus was inactivated by treatment at 56°C for 1 h. The final

stock contained 6.8×10^6 genomes per μl as determined by Southern analysis (33). The level of contaminating wild-type AAV was $<2.3 \times 10^4$ genomes per μl as determined by Southern analysis using AAV coding sequences as a probe.

Plasmid constructions. To construct pASNori2 (see Fig. 1), a *Bam*HI-*Esp*3I fragment of pSV2neo (41) containing an *Ssp*I-*Bst*1107I origin fragment from pACYC184 (7) in the *Bst*BI site (end filled with the Klenow fragment of DNA polymerase I) downstream of the *neo* (neomycin phosphotransferase) gene was inserted in the *Bgl*II sites of the AAV vector backbone of pTR (34) after attachment of *Bam*HI linkers to the pSV2neo *Esp*3I site. The pACYC184 fragment contains the p15A bacterial plasmid origin (10), with the direction of leading-strand DNA synthesis opposite that of *neo* gene transcription. pASNori2 also contains the pMB1 plasmid origin (5) from pBR322 (6). pASNori1, a deletion derivative of pASNori2 lacking the pMB1 origin, was constructed by end filling and circularizing a *Bsa*I-*Bst*1107I fragment of pASNori2.

Isolation of HeLa genomic DNA. Two confluent 10-cm-diameter dishes of each HeLa clone (2×10^7 cells) were lysed in genomic DNA lysis buffer (10 mM Tris [pH 8], 1 mM EDTA, 200 mM NaCl, 0.5% sodium dodecyl sulfate, 200 μg of proteinase K per ml) at 37°C overnight. The samples were then extracted with phenol and chloroform, extracted with butanol, and precipitated with 2 volumes of ethanol overnight at −20°C. The DNA was pelleted at 6,000 rpm in a Sorvall HS-4 rotor at 4°C for 25 min, washed with 70% ethanol, and air dried briefly. The DNA was resuspended in 500 μl of TE (10 mM Tris [pH 8], 1 mM EDTA), digested with 10 μg of RNase A (Sigma, St. Louis, Mo.) at 37°C for 3 h, extracted with phenol and chloroform, and precipitated with 50 μl of 3 M sodium acetate and 1 ml of ethanol at −20°C. Each pellet was resuspended in 200 μl of TE.

Provirus recovery in bacteria. To recover integrated AAV-SNori proviruses, 10 μg of HeLa genomic DNA was treated with 20 U of calf intestinal phosphatase (Boehringer Mannheim, Indianapolis, Ind.) to prevent ligation of free ends in the sample, heat inactivated at 65°C for 1 h, extracted with phenol and chloroform, and precipitated with ethanol. The resuspended DNA was digested with 20 U of *Eco*RI, which does not cut in the SNori vector, at 37°C for 4 h, heat inactivated at 65°C for 30 min, extracted with phenol and chloroform, and precipitated with ethanol. The resulting DNA fragments were resuspended and circularized with 200 U of T4 DNA ligase in 400 μl at 14°C overnight. The DNA was precipitated and one-fifth of the sample (approximately 2 μg) was electroporated into supercompetent *Escherichia coli* XL1Blue MRF′ cells (Stratagene, La Jolla, Calif.).

Provirus excision and amplification. HeLa cells and each of the nine HeLa clones from which plasmid had been recovered (see Table 1) were plated at 10^6 cells per 35-mm-diameter dish (Corning, Corning, N.Y.). The next day, cells were infected with wild-type AAV type 2 at 10 replication-competent particles/cell and adenovirus type 5 at 10 PFU/cell or were left uninfected as indicated in Fig. 6. Forty-four hours after infection, episomal DNA was isolated by the method of Hirt (21), with an additional proteinase K digestion, extraction with phenol and chloroform, and precipitation with ethanol. One-fourth of the episomal DNA from each dish was separated by alkaline agarose gel electrophoresis and transferred to Hybond-N+ membranes (Amersham, Arlington Heights, Ill.) according to standard procedures (35). Standards were prepared by *Bsm*I digest of pASNori2 to produce a 2.6-kb fragment. DNA was detected by Southern analysis using a *neo* gene probe.

DNA techniques. Restriction enzymes, T4 DNA ligase, and DNA polymerases were from New England BioLabs, Beverly, Mass. Proteinase K was from Boehringer Mannheim. Enzyme reactions were performed by using the manufacturer's recommended conditions. DNA manipulation and Southern blot analysis were performed by standard procedures (35). Southern blots were quantitated using a PhosphorImager 400S (Molecular Dynamics, Sunnyvale, Calif.). Plasmids were prepared by using Qiagen (Chatsworth, Calif.) columns. Dye terminator cycle sequencing was carried out with an AmpliTaq FS polymerase sequencing kit (Perkin-Elmer, Foster City, Calif.) and analyzed on an Applied Biosystems Inc. (Foster City, Calif.) sequencer. Oligonucleotides for sequencing are P1 (5′-dTACAAATAAAGCAATAGCATCAC-3′) and P2 (5′-dCCTCTGACAC ATGCAGCTC-3′). Sequences were analyzed by the Wisconsin version of the Genetics Computer Group program, using FASTA with default parameters and searching against GenBank/EMBL. HeLa flanking sequences were compared with the human *Alu* repetitive sequence humalurp7 (GenBank accession no. M57427), using FASTA with default parameters; a sequence with more than 70% identity over at least 100 nucleotides was considered a match. The probe used for detecting vector genomes by Southern analysis contained internal AAV-SNori sequences including the *neo* gene. The probe used for detecting site-specific integration was a 1.7-kb *Eco*RI-*Bam*HI fragment from the chromosome 19 integration locus cloned in pRE2 (38).

RESULTS

Transduction of HeLa cells by AAV-SNori. We designed an AAV shuttle vector (AAV-SNori) that could be recovered as a bacterial plasmid along with flanking human DNA after integration into host chromosomes. The recovered proviral plasmid could then be propagated in bacteria, and the junction fragments could be sequenced. The AAV-SNori vector contains the *neo* gene under the control of both the simian virus 40 (SV40) early promoter and the transposon 5 (Tn*5*) promoter for expression in human or bacterial cells, as well as the p15A bacterial replication origin packaged between AAV2 TRs (Fig. 1). These elements can support replication and confer kanamycin resistance in *E. coli*. The p15A origin was chosen because it replicates at the relatively low rate of 12 to 15 copies per *E. coli* chromosome (10), and we have found that plasmids containing the AAV TRs are more stable if they replicate at lower copy numbers (data not shown).

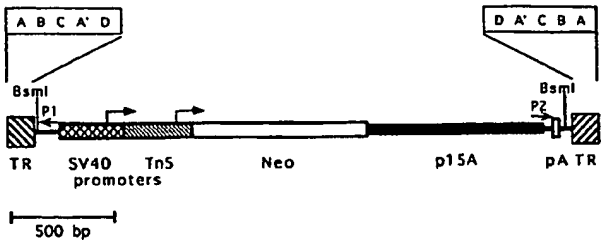

FIG. 1. AAV-SNori shuttle vector. The structure of the packaged genome of the AAV-SNori vector is shown; features include the neomycin phosphotransferase gene (Neo), the SV40 and Tn*5* promoters with their transcription start sites (arrows), the p15A replication origin, the AAV TRs, and the polyadenylation site (pA). An expanded diagram of the TR structure is shown to indicate the repeat domains. P1 and P2 represent binding sites for sequencing primers. The positions of *Bsm*I restriction sites are shown.

HeLa cervical carcinoma cells were transduced with purified AAV-SNori. One day after infection, the cells were treated with trypsin, and dilutions were plated in culture medium without selection. Selection was begun 24 h later with the antibiotic G418. Approximately 3×10^4 intact vector particles were required to produce a single, stable, G418-resistant colony. As all of these colonies contain integrated proviruses (see below), this represents the minimal vector integration frequency. The transient transduction rate was higher, as many of the colonies visible at earlier time points did not survive continued selection in G418, perhaps due to episomal vector gene expression. Fourteen G418-resistant colonies were isolated and expanded to approximately 2×10^7 cells under selection. Genomic DNA was isolated from each of 14 clones, digested with *Bsm*I or *Eco*RI, and analyzed by Southern blots probed with the *neo* gene (Fig. 2). *Bsm*I digests inside each of the TRs, so intact, integrated vector genomes should contain a 2.6-kb *Bsm*I fragment (Fig. 1). Figure 2A shows that 10 of the 14 clones had a fragment of the expected size. Clones 2, 6, 8, and 9 had deletions or rearrangements. The high-molecular-weight vector band in clone 8 DNA is faint, presumably due to a partial deletion of vector sequences. Clone 10 had a second, smaller fragment containing at least a portion of the *neo* gene. Quantitative analysis of the blots was consistent with single-copy integrated proviruses for all the clones except 8, 12, and 14, with the latter two clones containing four to five copies/cell.

Digestion of the transduced HeLa clones with *Eco*RI produced bands of various sizes ranging from 4 to 23 kb (Fig. 2B). As the AAV-SNori vector genome does not contain an *Eco*RI site, this finding is consistent with random integration. Although two *Bsm*I vector fragments were present in the DNA from clone 10 (Fig. 2A), only a single *Eco*RI fragment was detected, suggesting that both *Bsm*I fragments were located at a common integration site.

Provirus recovery as plasmids. Integrated vector proviruses were recovered from the transduced HeLa clones by digestion with *Eco*RI, circularization with DNA ligase, and transfer into *E. coli*. In this strategy, the *Eco*RI fragments containing intact

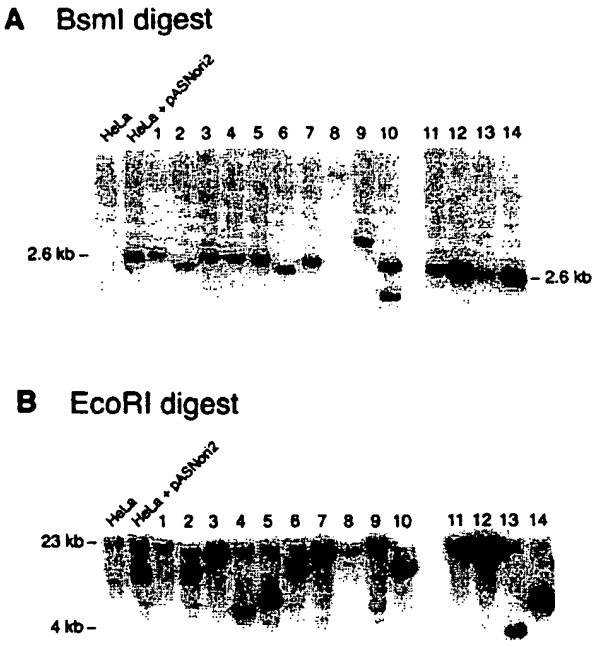

FIG. 2. Southern blots of transduced HeLa clones. Ten micrograms of genomic DNA from each HeLa clone (lanes 1 through 14) was digested with either *Bsm*I (A) or *Eco*RI (B), electrophoresed through a 1.2% agarose gel, transferred to a nylon membrane, and probed with internal vector sequences to detect integrated AAV-SNori proviruses. Lane HeLa, 10 μg of nontransduced HeLa DNA; lane HeLa + pASNori2, 10 μg of nontransduced HeLa DNA with 10 pg of pASNori2 plasmid DNA. This image was prepared by using Adobe Photoshop (Mountain View, Calif.) software.

vector proviruses also contain flanking chromosomal DNA, and their expected sizes can be predicted from the Southern analysis in Fig. 2B. Ligation at low DNA concentration enhances circularization of individual *Eco*RI fragments. The circularized fragments containing vector proviruses are then propagated as bacterial plasmids after electroporation of bacteria and selection with kanamycin. We recovered proviral plasmids from 9 of 14 HeLa transductants. Plasmids could not be recovered from the five other HeLa clones, even after repeated transformations (Table 1). This could be due to deletion or rearrangement of a region essential for plasmid function in bacteria but not required for G418 resistance in mammalian cells, such as the p15A origin or Tn*5* promoter. Four of the five HeLa clones containing proviruses that could not be recovered had an altered internal vector *Bsm*I fragment by Southern analysis (Fig. 2A), consistent with vector rearrangements.

Recovered plasmids were isolated from several of the kanamycin-resistant bacterial colonies and digested with *Eco*RI (Table 1). Most of these contained a single *Eco*RI site and were the size expected from Southern analysis (Fig. 2B). Transformation efficiencies varied considerably and did not appear to correlate with plasmid size. Only those plasmids with a single *Eco*RI fragment of the predicted size were considered correct. Of the plasmids that did not appear correct, 12 of 13 contained one or more extra *Eco*RI fragments in addition to the predicted fragment, which were presumably acquired by intermolecular ligation. Multiple proviral *Eco*RI fragments recovered from each cell line had the same restriction pattern, suggesting that major DNA rearrangements had not occurred

TABLE 1. Recovered provirus plasmids[a]

HeLa clone	Size of *Bsm*I fragment (kb)	Size of *Eco*RI fragment (kb)	Transformation efficiency/μg of DNA	No. of plasmids analyzed	No. of plasmids correct
1	2.6	19	<0.17	0	0
2	2.3	8	<0.17	0	0
3	2.6	16	2.5	5	5
4	2.6	4.7	101	5	5
5	2.6	5.5	68	5	4
6	2.3	7	<0.17	0	0
7	2.6	17	18	5	3
8	6.5	11	<0.17	0	0
9	3.3	16	<0.12	0	0
10	2.6, 1.9	9	2	4	2
11	2.6	20	9.5	5	5
12	2.6	18	4	8	1
13	2.6	4	1.5	3	3
14	2.6	5.5	4.5	5	4

[a] Integrated proviruses were recovered from transduced HeLa clones and transformed into bacteria. The sizes of fragments from Southern blots (Fig. 2), the transformation efficiency, and the number of plasmids that had a single *Eco*RI fragment (no. of plasmids correct) are listed for each HeLa clone.

during the transfer to bacteria. Further restriction digests confirmed that each recovered plasmid contained a single vector provirus copy (data not shown).

Because the TR sequences might be unstable during provirus recovery and replication in bacteria, we performed control transformations to ensure that our assay could recover plasmids with intact TRs. Plasmid pASNori1 contains the entire AAV-SNori vector genome including both TRs and depends on the p15A replication origin in the vector sequences for replication, making it similar in structure to the recovered plasmids. pASNori1 was linearized with *Eco*RI, gel purified, mixed with nontransduced, *Eco*RI-digested HeLa DNA, and subjected to the same recovery procedure as integrated proviruses. Of 10 recovered plasmids, 8 contained the expected unique *Eco*RI site. Two recovered plasmids contained extra *Eco*RI fragments in addition to the expected pASNori1 fragment. The restriction digest patterns of all eight correct plasmids were identical to the original pASNori1 plasmid after digestion with four restriction enzymes that have sites in the TRs, confirming that the TRs remained intact during our recovery procedure (data not shown).

Sequence analysis of integration junctions and flanking genomic DNA. One example of each correct recovered plasmid was sequenced by using primers P1 and P2 (Fig. 1). These primers bind to locations in the vector genome about 70 nucleotides inside each of the TRs. The sequenced regions begin at internal vector sequences and then proceed outward through the TR region, the integration junction site, and flanking human DNA. Of the nine plasmids analyzed, flanking human sequence information of 400 or more nucleotides was obtained from 14 of the 18 possible integration junctions. Three other junctions (12L, 12R, and 14L) contained nearly complete TRs that stalled the sequencing polymerase and one (10R) contained TR sequences joined to Tn*5* sequences (see below). The sequences around each junction site are shown in Fig. 3 and are designated as from the left TR (L) or right TR (R) of each recovered provirus. Figure 3A displays junction sequences that could be aligned in the "flip" orientation; Fig. 3B shows a right junction sequence that was in the "flop" orientation. In the flop orientation, the B and C regions are inverted compared to the flip orientation (D-A′-B-C-A instead of D-A′-C-B-A) (29). Figure 3C shows a left junction sequence

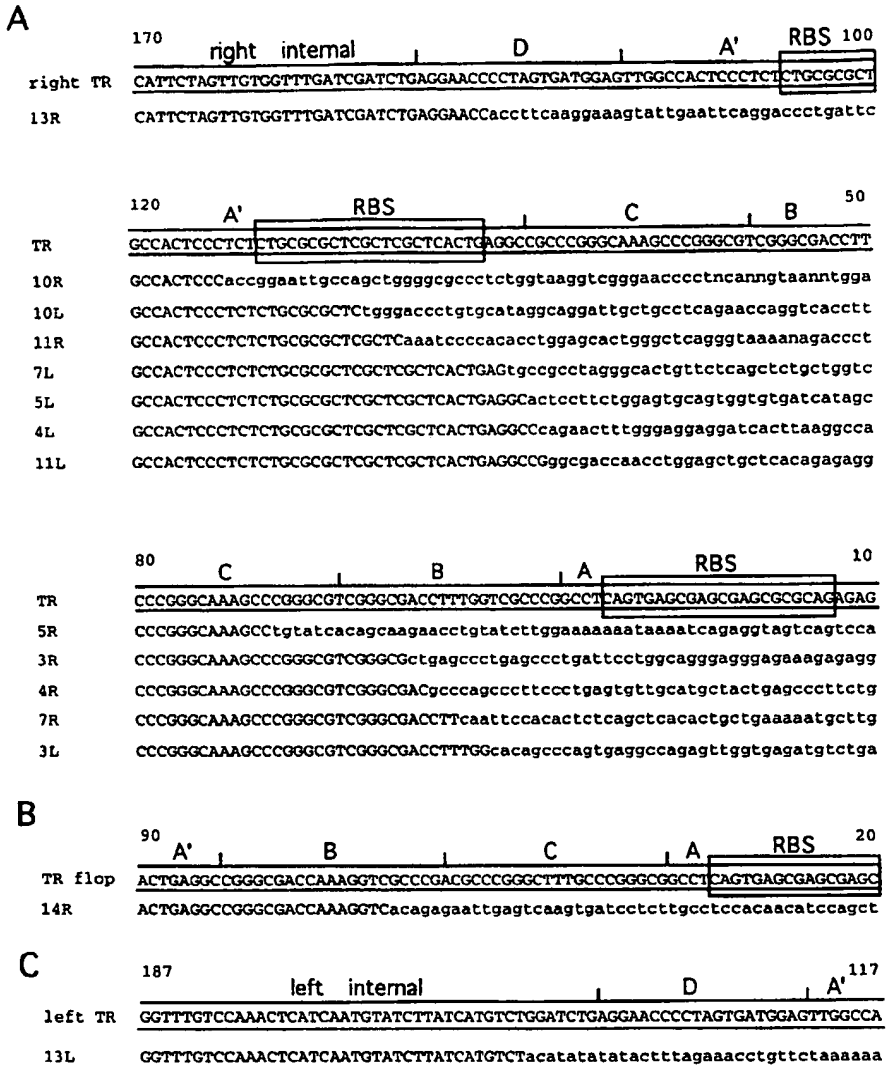

FIG. 3. AAV vector integration junctions. Portions of AAV vector sequences are shown in uppercase. Up to 62 nucleotides of flanking sequences are shown for each junction in lowercase and are aligned beneath the intact vector sequence (TR). Nucleotide positions are numbered beginning at the end of the intact TR. The sequences are designated as from the left (L) or right (R) of each numbered provirus clone. The Rep binding site (RBS) (9, 34) is boxed. Subdomains of the TRs (A, B, C, A', and D) are indicated above the sequences. The flanking sequence is HeLa DNA, except for junction 10R, where it is Tn*5* DNA (see text). (A) Sequences aligned to the right TR in the flip orientation and internal vector DNA. (B) Sequence aligned to the flop orientation of the TR. (C) Sequence aligned to the left internal portion of the vector.

with the recombination site internal to the TR. Of the six junctions containing part of the B or C region in which the orientation could be determined, five were in the flip orientation.

The sequence data show that no two proviruses had the same junction site. Junctions were dispersed throughout the TRs and internal vector sequences. We could not identify a common sequence motif in vector or chromosomal DNA at the junction sites. None of the flanking sequences were similar to each other, and there were no matches with any gene in the GenBank or EMBL database. The only known sequence element found was a single *Alu* repeat in the flanking human DNA of junction 4R.

Comparison of the flanking sequences with two sequences from the chromosome 19 integration locus, AAVS1 (GenBank accession no. S51329) and pRE2 (provided by R. J. Samulski), produced no matches, indicating that the clones did not integrate into the site-specific integration locus of wild-type AAV2. This observation was confirmed by probing Southern blots of HeLa clone DNAs with a portion of the site-specific integration locus, which showed no rearrangements at this locus in any of the 14 transductants analyzed (data not shown). In addition, analysis of the flanking sequences revealed no binding sites for the AAV Rep protein (9, 34), in contrast to the chromosomal site-specific integration locus (16).

Two of the junctions (4L and 5R) had structures that were the result of rearrangements. A schematic representation of these proviruses is shown in Fig. 4. The rearrangements in sequences 4L and 5R were in the regions just inside each TR where a 44-nucleotide region is repeated in the left and right

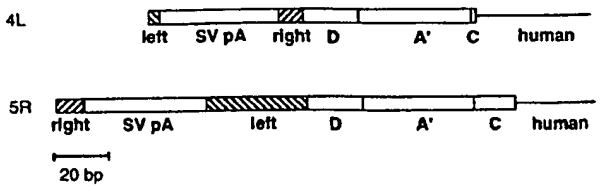

FIG. 4. Structures of rearranged sequences. The 4L and 5R provirus junction sequences are diagrammed to show where internal vector sequences recombined. The positions of left and right internal vector sequences, the common SV40 polyadenylation site found in either end (SV pA), the TR domains (D, A', and C), and flanking human DNA are shown.

portions of the vector. This region contains the SV40 polyadenylation signal and was duplicated as a result of vector construction. In the case of junction 4L, sequencing from the left P1 primer (Fig. 1) read a portion of left internal sequences, the common SV40 sequence, and then right internal vector sequences followed by a portion of the right TR and flanking chromosomal DNA. A similar finding was observed in junction 5R reading from the right P2 primer. Apparently homologous recombination occurred between these sites and produced an inverted internal portion of the vector. Recombination could have happened during the production of the AAV-SNori stocks or the transduction of HeLa cells.

Junction (10R) was not joined to human DNA. This junction contained the A' region of the right TR joined to a second copy of vector Tn*5* DNA. The sequence around the junction site is shown in Fig. 3. This explains the two *neo*-hybridizing bands detected in Fig. 2A for HeLa clone 10. Flanking HeLa DNA sequence was not obtained for this junction and must lie beyond the region sequenced.

When the TR is drawn in its predicted secondary structure, the recombination sites of the integrated proviruses are scattered throughout the repeat (Fig. 5). Except for some potential clustering of recombination events between the C and A' repeat domains, our data set does not suggest that secondary structure in the TRs directs a specific integration process. Additional junctions will need to be sequenced to determine if the four junctions observed between C and A' represent a recombination hotspot.

Flanking DNA sequence was not obtained for the remaining 3 of the 18 possible integration junctions recovered (12L, 12R, and 14L). The sequences obtained for these junctions ended abruptly in the A' region of the TR, where the sequencing polymerase apparently stalled. Based on our sequencing of intact AAV TRs, this is consistently observed when the secondary structure of the inverted repeat remains intact, and the polymerase stalls at the first base that can pair in the secondary structure (data not shown). If we assume that this also occurred while we were sequencing these three junctions, then we can place the end of the vector provirus DNA at the last base of the predicted secondary structure that would have stalled the sequencing polymerase, in the corresponding positions of the A domain. These junctions are indicated with asterisks in Fig. 5.

Provirus rescue by wild-type infection. Integrated wild-type AAV proviruses can be excised and amplified by infection with adenovirus (3, 8, 18, 26), and a similar rescue of vector proviruses can occur by infection with wild-type AAV and adenovirus (19, 30, 36, 42). We screened each of the nine transduced HeLa clones that contained proviruses recovered in bacteria to see if their TR structures would correlate with the potential for excision and amplification. Only clones 12 and 14 contained proviruses that could be rescued by coinfection with wild-type AAV and adenovirus, and none of the clones could be rescued by adenovirus alone, as expected if no *rep* gene was present in the cells. Figure 6 shows representative results from this experiment for clones 12, 13, and 14. The two rescuable proviruses had the most intact TRs of all the proviruses that we analyzed (Fig. 5), and the nearly complete repeats on both ends of clone 12 correlated with more efficient rescue than those of clone 14.

DISCUSSION

We have developed a shuttle vector system to study the structure of integrated AAV vector proviruses and analyze their flanking DNA. Fourteen independent, G418-resistant HeLa clones were isolated after transduction with an AAV vector containing the *neo* gene, and integrated vector proviruses were recovered as bacterial plasmids from nine of these lines. Sequence analysis of the proviral junction sites showed that each integration event occurred at a different chromosomal site, and none of these sites were at the chromosome 19

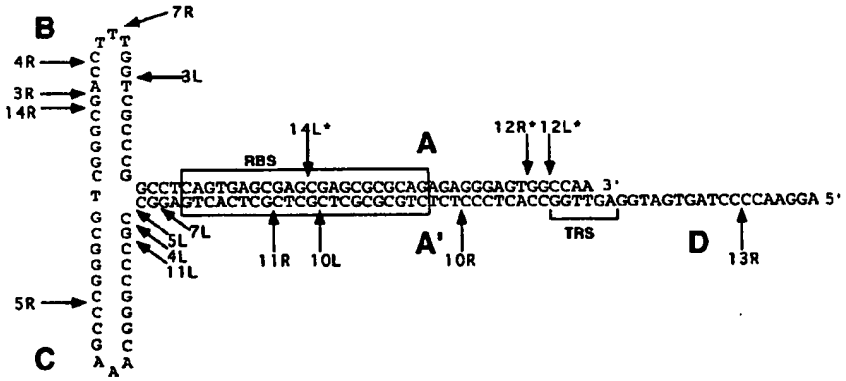

FIG. 5. Junction sites and TR secondary structure. Junction recombination sites within the TRs of recovered proviruses are shown in relation to the secondary structure of the TR. All sequences were adapted to be displayed in the flip orientation. Junctions 12L, 12R, and 14L are presumed junction sites because the actual sequences ended in the corresponding positions in the A' region, suggesting that the TR was intact up to that point. The positions of the Rep binding site (RBS) (9, 34) and terminal resolution site (TRS) (40) are shown.

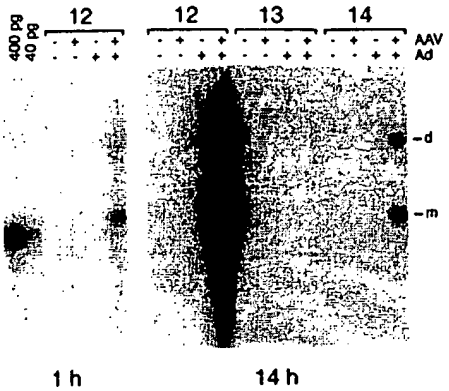

FIG. 6. Provirus rescue. Transduced HeLa clones were infected with wild-type AAV2 (AAV), and/or adenovirus (Ad) or left uninfected as indicated. Episomal DNA was isolated from the cells 44 h later, electrophoresed through an alkaline agarose gel, and subjected to Southern analysis using a *neo* gene probe. The positions of the monomer (m) and dimer replicative (d) vector forms are indicated. The standards are 400 and 40 pg of the 2.6-kb *Bsm*I fragment of pASNori2. The results for HeLa clones 12 to 14 are shown (14-h exposure), with HeLa clone 12 also shown at a shorter exposure (1 h). This image was prepared by using Adobe Photoshop software.

site-specific integration locus of wild-type AAV. Each recombination event was at a different position in the vector genome, and no fully intact vector proviruses were recovered. There was no sequence homology between the vector and flanking human DNA and none between the different flanking sequences recovered. These results demonstrate that AAV vector integration occurs at random chromosomal locations and that each integrated provirus contains variable amounts of the complete vector genome.

The recovered proviruses were joined to human DNA sequences at nucleotide positions throughout the vector TRs and at internal vector sequences in one case. Because our system required the presence of functional promoters, *neo* gene, and bacterial origin, the ends of the vector were the only portions that could have been disrupted by integration and still produced a functional provirus capable of being recovered. Southern analysis of the proviruses that could not be recovered showed that four of five contained rearrangements within internal vector sequences, suggesting that the junctions occurred inside the *Bsm*I sites (Fig. 2A). While our results are consistent with random recombination throughout the vector genome, we cannot exclude a partial preference for TR recombination sites, since more junctions should have been observed in the approximately 170 nonessential internal vector nucleotides if a completely random representation had been recovered.

Based on Southern analysis and restriction mapping of recovered plasmids, 11 of 14 HeLa transductants contained one integrated vector genome per cell, and clone 10 contained a partially duplicated vector genome. Two clones (12 and 14) had a stronger hybridization signal consistent with four to five vector copies per cell. In these clones, the vector signal was present in unique bands after digestion outside of vector DNA with *Eco*RI, and their recovered plasmids contained one copy of the vector provirus within the *Eco*RI fragment. This finding suggests that amplification of the provirus and flanking DNA contained in the *Eco*RI fragment occurred after integration of the vector, perhaps during the selection process in G418. It is possible that proviral sequences were responsible for this amplification process, as the AAV TR has been shown to function as a replication origin in carcinogen-treated mammalian cells (46), and these two proviruses contained the most complete TR structures of all those recovered.

Both concatemeric and single-copy vector proviruses have been described in prior studies of vector integration based on Southern analysis of transduced, cultured cells (11, 27, 30, 33, 36, 44). More recently, animal studies have demonstrated the presence of concatemeric vector sequences in high-molecular-weight DNA by PCR after in vivo vector administration (14, 20, 45). The conditions favoring concatemeric versus single-copy integration are unknown and could include cell-type-specific functions, multiplicity of infection, contaminating *rep*$^+$ helper virus particles, and differences in selection conditions. None of the G418-resistant HeLa cell transductants that we analyzed contained concatemeric vector proviruses, over a range of infection multiplicities of 0.17 to 170 vector particles/cell, with $\leq$0.3% wild-type AAV contamination (wild-type genomes per vector genome). Our results are in agreement with previous studies suggesting that concatemeric vector proviruses are preferentially rescued by coinfection with wild-type AAV and adenovirus (30, 36), as most of the integrants that we analyzed could not be rescued. Although the two transductants that could be rescued were not present as concatemers joined at the TRs, they were present at multiple copies per cell. These proviruses also had the smallest deletions in their TRs, with at least one repeat containing a complete copy of each repeat domain sequence. This proviral structure would be expected to regenerate a fully intact TR by gene conversion as previously shown for wild-type AAV (37).

There were both differences and similarities between vector integration in our system and integration by wild-type AAV2. Unlike the random vector integration that we observed, two-thirds of wild-type proviruses studied in latently infected human cells were found in the site-specific integration locus on chromosome 19 (25, 38). Although these wild-type integration events occurred at a common locus, the junctions were at different nucleotide positions within the locus and at different positions in the AAV TRs (23, 38). This heterogeneity in junction sites within the TRs is similar to our findings with vector integration. Site-specific integration into chromosome 19 requires the presence of a Rep binding site in the chromosomal target DNA (16, 28), and no Rep binding sites were identified in the human flanking sequences that we recovered. This finding suggests that vector integration is not mediated by the Rep protein, including Rep protein molecules that might be contained in the entering vector particle. Presumably, either Rep expression in the transduced cell is necessary for site-specific integration or wild-type virions contain functional Rep molecules that are absent from vector virions. A related observation may be that integrated wild-type proviruses are often found as concatemers (8, 22, 26), which we did not observe for vector proviruses. This observation supports the hypothesis that site-specific integration involves Rep-dependent replication of the chromosome and virus during integration (28, 43), while Rep-independent vector integration occurs by another, nonreplicative mechanism.

Our data suggest that AAV vector integration is a nonhomologous recombination reaction. While no clear recombination hotspots were identified in the vector genome, there may have been a clustering of junctions between the C and A′ regions of the TR, suggesting a possible role for Holliday-like recombination intermediates in the reaction. Although the exact mechanism has not been determined, the recombination event is presumably mediated by cellular proteins. Cleavage of the chromosomal preintegration site may be due to host nucleases or DNA damage. The requirement for chromosomal breaks could in part explain why DNA-damaging agents in-

crease transduction by AAV vectors (2, 32). The lack of specific vector recombination enzymes also suggests why vector integration is such an inefficient process, requiring hundreds to thousands of vector particles per integration event (15, 19, 33, 36). One explanation for the consistent recovery of incomplete proviruses is that the entering, linear vector genomes are partially degraded before the integration event. Despite our determination of the structure of integrated proviruses, key steps in the recombination process have yet to be elucidated, including the role of vector second-strand synthesis, which is also associated with increased transduction rates (12, 13), and whether the recombination reaction involves single-stranded or double-stranded vector molecules. The nature of the chromosomal preintegration site is completely unknown, and there could be large deletions or even interchromosomal crossover events associated with vector integration.

The uncertain nature of the AAV vector integration process has important consequences for gene therapy. In contrast to retroviral vectors, which integrate in a precise and predictable reaction mediated by the viral integrase protein, each AAV vector integration event is likely to produce a different, incomplete provirus, with unknown effects on the host chromosomes involved. Many of the integrated vectors could contain deletions of the transgene being delivered, with resulting decreases in transduction efficiencies. As our understanding of the process improves, methods for increasing the efficiency and predictability of vector integration that improve the prospects for gene therapy by AAV vectors may be developed.

ACKNOWLEDGMENTS

We thank Roli Hirata and Jaclynn Mac for expert technical assistance.

This work was supported by grants from the Cystic Fibrosis Foundation, the American Society of Hematology, and the National Institutes of Health (P01 HL53750 and K08 HL03100).

REFERENCES

1. **Afione, S. A., C. K. Conrad, W. G. Kearns, S. Chunduru, R. Adams, T. C. Reynolds, W. B. Guggino, G. R. Cutting, B. J. Carter, and T. R. Flotte.** 1996. In vivo model of adeno-associated virus vector persistence and rescue. J. Virol. 70:3235–3241.
2. **Alexander, I. E., D. W. Russell, and A. D. Miller.** 1994. DNA-damaging agents greatly increase the transduction of nondividing cells by adeno-associated virus vectors. J. Virol. 68:8282–8287.
3. **Berns, K. I., T. C. Pinkerton, G. F. Thomas, and M. D. Hoggan.** 1975. Detection of adeno-associated virus (AAV)-specific nucleotide sequences in DNA isolated from latently infected Detroit 6 cells. Virology 68:556–560.
4. **Bertran, J., J. L. Miller, Y. Yang, A. Fenimore-Justman, F. Rueda, E. F. Vanin, and A. W. Nienhuis.** 1996. Recombinant adeno-associated virus-mediated high-efficiency, transient expression of the murine cationic amino acid transporter (ecotropic retroviral receptor) permits stable transduction of human HeLa cells by ecotropic retroviral vectors. J. Virol. 70:6759–6766.
5. **Betlach, M., V. Hershfield, L. Chow, W. Brown, H. Goodman, and H. W. Boyer.** 1976. A restriction endonuclease analysis of the bacterial plasmid controlling the EcoRI restriction and modification of DNA. Fed. Proc. 35:2037–2043.
6. **Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, and H. W. Boyer.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene 2:95–113.
7. **Chang, A. C., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. 134:1141–1156.
8. **Cheung, A. K., M. D. Hoggan, W. W. Hauswirth, and K. I. Berns.** 1980. Integration of the adeno-associated virus genome into cellular DNA in latently infected human Detroit 6 cells. J. Virol. 33:739–748.
9. **Chiorini, J. A., S. M. Wiener, R. A. Owens, S. R. Kyöstiö, R. M. Kotin, and B. Safer.** 1994. Sequence requirements for stable binding and function of Rep68 on the adeno-associated virus type 2 inverted terminal repeats. J. Virol. 68:7448–7457.
10. **Cozzarelli, N. R., R. B. Kelly, and A. Kornberg.** 1968. A minute circular DNA from Escherichia coli 15. Proc. Natl. Acad. Sci. USA 60:992–999.
11. **Einerhand, M. P., M. Antoniou, S. Zolotukhin, N. Muzyczka, K. I. Berns, F. Grosveld, and D. Valerio.** 1995. Regulated high-level human beta-globin gene expression in erythroid cells following recombinant adeno-associated virus-mediated gene transfer. Gene Ther. 2:336–343.
12. **Ferrari, F. K., T. Samulski, T. Shenk, and R. J. Samulski.** 1996. Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. J. Virol. 70:3227–3234.
13. **Fisher, K. J., G. P. Gao, M. D. Weitzman, R. DeMatteo, J. F. Burda, and J. M. Wilson.** 1996. Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. J. Virol. 70:520–532.
14. **Fisher, K. J., K. Jooss, J. Alston, Y. Yang, S. E. Haecker, K. High, R. Pathak, S. E. Raper, and J. M. Wilson.** 1997. Recombinant adeno-associated virus for muscle directed gene therapy. Nat. Med. 3:306–312.
15. **Flotte, T. R., R. Solow, R. A. Owens, S. Afione, P. L. Zeitlin, and B. J. Carter.** 1992. Gene expression from adeno-associated virus vectors in airway epithelial cells. Am. J. Respir. Cell. Mol. Biol. 7:349–356.
16. **Giraud, C., E. Winocour, and K. I. Berns.** 1995. Recombinant junctions formed by site-specific integration of adeno-associated virus into an episome. J. Virol. 69:6917–6924.
17. **Graham, F. L., J. Smiley, W. C. Russell, and R. Nairn.** 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J. Gen. Virol. 36:59–74.
18. **Handa, H., K. Shiroki, and H. Shimojo.** 1977. Establishment and characterization of KB cell lines latently infected with adeno-associated virus type 1. Virology 82:84–92.
19. **Hermonat, P. L., and N. Muzyczka.** 1984. Use of adeno-associated virus as a mammalian DNA cloning vector: transduction of neomycin resistance into mammalian tissue culture cells. Proc. Natl. Acad. Sci. USA 81:6466–6470.
20. **Herzog, R. W., J. N. Hagstrom, S. H. Kung, S. J. Tai, J. M. Wilson, K. J. Fisher, and K. A. High.** 1997. Stable gene transfer and expression of human blood coagulation factor IX after intramuscular injection of recombinant adeno-associated virus. Proc. Natl. Acad. Sci. USA 94:5804–5809.
21. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. 26:365–369.
22. **Kotin, R. M., and K. I. Berns.** 1989. Organization of adeno-associated virus DNA in latently infected Detroit 6 cells. Virology 170:460–467.
23. **Kotin, R. M., R. M. Linden, and K. I. Berns.** 1992. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. EMBO J. 11:5071–5078.
24. **Kotin, R. M., J. C. Menninger, D. C. Ward, and K. I. Berns.** 1991. Mapping and direct visualization of a region-specific viral DNA integration site on chromosome 19q13-qter. Genomics 10:831–834.
25. **Kotin, R. M., M. Siniscalco, R. J. Samulski, X. D. Zhu, L. Hunter, C. A. Laughlin, S. McLaughlin, N. Muzyczka, M. Rocchi, and K. I. Berns.** 1990. Site-specific integration by adeno-associated virus. Proc. Natl. Acad. Sci. USA 87:2211–2215.
26. **Laughlin, C. A., C. B. Cardellichio, and H. C. Coon.** 1986. Latent infection of KB cells with adeno-associated virus type 2. J. Virol. 60:515–524.
27. **Lebkowski, J. S., M. M. McNally, T. B. Okarma, and L. B. Lerch.** 1988. Adeno-associated virus: a vector system for efficient introduction and integration of DNA into a variety of mammalian cell types. Mol. Cell. Biol. 8:3988–3996.
28. **Linden, R. M., E. Winocour, and K. I. Berns.** 1996. The recombination signals for adeno-associated virus site-specific integration. Proc. Natl. Acad. Sci. USA 93:7966–7972.
29. **Lusby, E., K. H. Fife, and K. I. Berns.** 1980. Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA. J. Virol. 34: 402–409.
30. **McLaughlin, S. K., P. Collis, P. L. Hermonat, and N. Muzyczka.** 1988. Adeno-associated virus general transduction vectors: analysis of proviral structures. J. Virol. 62:1963–1973.
31. **Muzyczka, N.** 1992. Use of adeno-associated virus as a general transduction vector for mammalian cells. Curr. Top. Microbiol. Immunol. 158:97–129.
32. **Russell, D. W., I. E. Alexander, and A. D. Miller.** 1995. DNA synthesis and topoisomerase inhibitors increase transduction by adeno-associated virus vectors. Proc. Natl. Acad. Sci. USA 92:5719–5723.
33. **Russell, D. W., A. D. Miller, and I. E. Alexander.** 1994. Adeno-associated virus vectors preferentially transduce cells in S phase. Proc. Natl. Acad. Sci. USA 91:8915–8919.
34. **Ryan, J. H., S. Zolotukhin, and N. Muzyczka.** 1996. Sequence requirements for binding of Rep68 to the adeno-associated virus terminal repeats. J. Virol. 70:1542–1553.
35. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
36. **Samulski, R. J., L. S. Chang, and T. Shenk.** 1989. Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. J. Virol. 63:3822–3828.
37. **Samulski, R. J., A. Srivastava, K. I. Berns, and N. Muzyczka.** 1983. Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats of AAV. Cell 33:135–143.
38. **Samulski, R. J., X. Zhu, X. Xiao, J. D. Brook, D. E. Housman, N. Epstein, and L. A. Hunter.** 1991. Targeted integration of adeno-associated virus

(AAV) into human chromosome 19. EMBO J. 10:3941–3950. (Erratum, 3:1228, 1992.)
39. **Scherer, W. F., J. T. Syverton, and G. O. Gey.** 1953. Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. J. Exp. Med. 97:695–709.
40. **Snyder, R. O., D. S. Im, T. Ni, X. Xiao, R. J. Samulski, and N. Muzyczka.** 1993. Features of the adeno-associated virus origin involved in substrate recognition by the viral Rep protein. J. Virol. 67:6096–6104.
41. **Southern, P. J., and P. Berg.** 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J. Mol. Appl. Genet. 1:327–341.
42. **Tratschin, J. D., I. L. Miller, M. G. Smith, and B. J. Carter.** 1985. Adeno-associated virus vector for high-frequency integration, expression, and rescue of genes in mammalian cells. Mol. Cell. Biol. 5:3251–3260.
43. **Urcelay, E., P. Ward, S. M. Wiener, B. Safer, and R. M. Kotin.** 1995. Asymmetric replication in vitro from a human sequence element is dependent on adeno-associated virus Rep protein. J Virol. 69:2038–2046.
44. **Walsh, C. E., J. M. Liu, X. Xiao, N. S. Young, A. W. Nienhuis, and R. J. Samulski.** 1992. Regulated high level expression of a human gamma-globin gene introduced into erythroid cells by an adeno-associated virus vector. Proc. Natl. Acad. Sci. USA 89:7257–7261.
45. **Xiao, X., J. Li, and R. J. Samulski.** 1996. Efficient long-term gene transfer into muscle tissue of immunocompetent mice by adeno-associated virus vector. J. Virol. 70:8098–8108.
46. **Yalkinoglu, A. O., H. Zentgraf, and U. Hubscher.** 1991. Origin of adeno-associated virus DNA replication is a target of carcinogen-inducible DNA amplification. J. Virol. 65:3175–3184.

JOURNAL OF VIROLOGY, Nov. 1995, p. 6917–6924
0022-538X/95/$04.00+0

Vol. 69, No. 11

Recombinant Junctions Formed by Site-Specific Integration of Adeno-Associated Virus into an Episome

CATHERINE GIRAUD,[1] ERNEST WINOCOUR,[2] AND KENNETH I. BERNS[1]*

Department of Microbiology, Hearst Microbiology Research Center, Cornell University Medical College, New York, New York 10021,[1] and Department of Molecular Genetics and Virology, Weizmann Institute of Science, Rehovot, Israel[2]

Received 5 June 1995/Accepted 14 August 1995

A model system using an episomal Epstein-Barr virus shuttle vector was recently developed to study the adeno-associated virus (AAV) site-specific integration event in chromosome 19q13.3-qter (C. Giraud, E. Winocour, and K. I. Berns, Proc. Natl. Acad. Sci. USA 91:10039–10043, 1994). In this study, we analyze the recombinant junctions generated after integration of the AAV genome into an Epstein-Barr virus shuttle vector carrying 8.2, 1.6, or 0.51 kb of the chromosome 19 preintegration sequence (AAVS1 locus). In most of the recombinants, one end of the viral genome was joined to a portion of the AAVS1 DNA previously shown to be a minimum target for AAV integration. Within this AAVS1 segment, the AAV insertion points were strikingly clustered around a binding site for the AAV regulatory protein. In all cases, the second junction with AAV occurred with vector DNA outside of the AAVS1 segment. With respect to the viral genome, one junction with the shuttle vector DNA occurred either within the AAV inverted terminal repeat (itr), or near the P5 promoter, approximately 100 nucleotides distal to a modified itr. The modified itr in 5 of 11 recombinants involved a head-to-tail organization. In one such instance, the AAV insert contained slightly more than one genome equivalent arranged in a head-to-tail manner with a junction close to the P5 promoter; the AAV insert in this recombinant episome could be rescued by adenovirus infection and replicated to virus particles. The significance of the head-to-tail organization is discussed in terms of the possible circularization of AAV DNA before or during integration.

Adeno-associated virus (AAV) has been recognized recently as a likely vector for gene delivery in mammalian cells (4, 21, 31, 43). One particularly attractive feature of this virus is its ability to integrate with a high degree of specificity in the AAVS1 locus (preintegration sequence) on chromosome 19q13.3-qter (23–25, 33, 36). By using an Epstein-Barr virus (EBV) shuttle vector (27), we have previously shown that an 8.2-kb AAVS1 preintegration DNA, propagated as an extrachromosomal episome, is a target for site-specific AAV DNA integration (12). Sequential deletion of the 8.2-kb sequence identified a minimum sequence of 510 nucleotides (nt) able to direct integration. Several signals potentially involved in the integration process were identified on this short piece of DNA: a Rep 78/68 binding site (6, 42), a potential terminal resolution site (38), and an M26 motif (32, 37). This 510-bp fragment was also associated with rearrangements of the shuttle vector genome (12).

To characterize the AAV site-specific integration event, we have analyzed the products of integration in EBV vectors carrying the AAVS1 preintegration locus. Sequencing of the junctions between AAV and the vector has shown that a highly prevalent (in 80% of the recombinants) insertion point is located close to the previously described Rep binding motif and terminal resolution sequence at the 5′ end of the AAVS1 preintegration locus. This finding provides further evidence underlining the importance of these AAV recognition signals in site-specific integration. In addition, the sequencing data have revealed that 5 of the 11 recombinant structures analyzed contain a head-to-tail junction.

* Corresponding author. Mailing address: Department of Microbiology, Hearst Microbiology Research Center, Cornell University Medical College, 1300 York Ave., New York, NY 10021. Phone: (212) 746-6505. Fax: (212) 746-8587. Electronic mail address: KBerns@mail.med.cornell.edu.

MATERIALS AND METHODS

Production and selection of EBV recombinant vectors. The C17 cell lines (3) propagating the p220.2 EBV shuttle vectors (9, 44) carrying segments of AAVS1 DNA, designated C17-p220.2(AAVS1 kb 0-8.2), C17-p220.2(AAVS1 kb 0-1.6), and C17-p220.2(AAVS1 kb 0-0.51), have been described previously (12). These cell lines were infected at several passage levels with AAV at an input multiplicity of 20 infectious units per cell. The extrachromosomal DNA was extracted 48 h postinfection with a slightly modified Hirt extraction procedure (18) and transfected into the SURE strain of *Escherichia coli* (14; Stratagene). The recombinants were isolated after colony hybridization of a replica filter with a nonradioactive single-stranded AAV genomic DNA probe as previously described (12).

Oligonucleotide probes and hybridization. AAV oligonucleotides (see Fig. 1) (AAV1, 5′-CTCACgTgACCTCTAATACAgg-3′; AAV2, 5′-gggACCTTAATC ACAATCTCg-3′; AAV3, 5′-CTTCTCggCCACggTCAgg-3′; AAV4, 5′-gAAAT gTCCTCCACgggCTg-3′; AAV5, 5′-CTTgTCgAgTCCgTTgAAgg-3′; AAV6, 5′-CATgAATCCTCTCATCgACC-3′; AAV7, 5′-gTggAgaTCgAgTgggAgC-3′; AA V8, 5′-ggCACCAgATACCTgACTCg-3′; AAV9, 5′-CTAgTTTCCATggCTAC g-3′; AAVD, 5′-ggAACCCCTAgTgATggAg-3′) and AAVS1 oligonucleotides (oligo 1, 5′-ACTTgCTAgTATgCCgTggg-3′; oligo2, 5′-CTACCTgCCCAgCAC ACC-3′; oligo 3, 5′-CATCCTCTCCggACATCg-3′), synthesized by the Oligo Etc Company, were labeled with the Dig-Oligonucleotide 3′ end Labeling Kit (Boehringer Mannheim). Vector DNA (100 ng) was dotted onto a nylon membrane (Hybond-N; Amersham), denatured with NaOH at 0.5 M plus NaCl at 1.5 M, neutralized with 0.5 M Tris-HCl (pH 8.0)–1.5 M NaCl, and UV cross-linked. The membranes were subsequently prehybridized for 2 h at 68°C and hybridized for 3 h at a temperature determined as optimum for each nucleotide (5 to 20°C under the melting temperature). Specific hybridization was detected with the Genius system kit (Boehringer Mannheim).

Sequencing reactions. Plasmid DNA was prepared by the Qiagen kit plasmid purification procedure (Qiagen Inc.), and sequencing reactions were performed with the Sequenase Quick-Denature plasmid sequencing kit (U.S. Biochemical). Oligonucleotides spread throughout the AAV genome (see Fig. 1) were used as primers for the sequencing reactions. To determine the junction sequences, an additional 15 primers (data available upon request) were utilized.

PCR for determination of head-to-tail organization of inverted terminal repeat (itr) and adjacent DNA. Plasmid DNA (10 ng) was mixed with 200 pmol each of oligonucleotides AAV1 and AAV8. The reaction was performed with the Boehringer PCR buffer (final concentrations, 10 mM Tris-HCl [pH 8.4], 50 mM

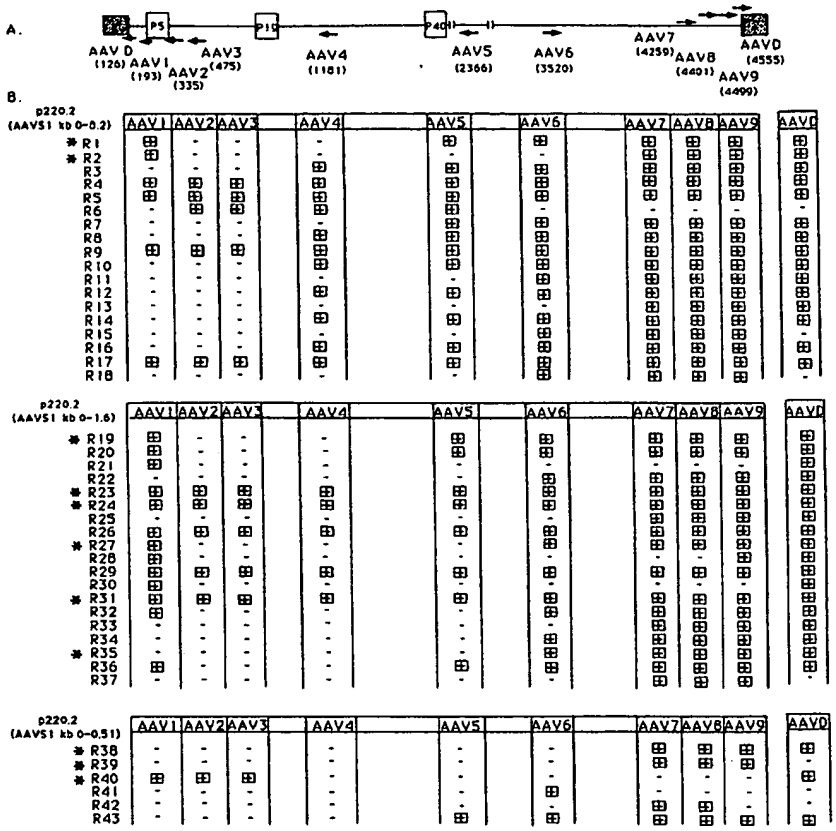

FIG. 1. Pattern of AAV sequences integrated in EBV recombinant vectors. (A) Schematic representation of the AAV genome showing the itr (▩), the promoters (▯), and the oligonucleotides (→) spanning the AAV internal sequences and the D region of the itr used as probes. The gap indicates a disruption in the sequence with respect to the scale of the drawing. (B) Summary of dot blot hybridizations between AAV oligonucleotides and EBV recombinant vectors (indicated by the letter R followed by a number) isolated from the C17-p220.2(AAVS1 kb 0-8.2), C17-p220.2(AAVS1 kb 0-1.6), and C17-p220.2(AAVS1 kb 0-0.51) cell lines 48 h postinfection with AAV. The asterisks show the recombinants which were subsequently sequenced at their junctions with the AAV genome (see Fig. 2 to 5).

KCl, 1.5 mM MgCl2, and 0.01% gelatin) supplemented by the four deoxynucleoside triphosphates at a final concentration of 100 each μM and 2.5 U of *Taq* DNA polymerase (Boehringer) in a total volume of 100 μl. Thirty cycles (denaturation at 94°C for 45 s, annealing at 58°C for 45 s, and elongation at 72°C for 1 min) were carried out with a Perkin Elmer Cetus (Norwalk, Conn.) PCR apparatus, and 10 μl of the reaction mixture was loaded onto a 1.5% agarose gel. Reactions were also performed under the same conditions but with only one of the primers.

Rescue of AAV from recombinant EBV shuttle vectors. C17 cells were infected with adenovirus type 2 (Ad2) at an input multiplicity of 10 infectious units per cell, for 1 h at 37°C and subsequently transfected with 20 μg of purified EBV recombinant plasmids by using the cationic lipid *N*-[1(2,3-dioleoyloxy)propyl)]-*N,N,N*-trimethylammonium methyl sulfate (DOTAP) (10, 39). Forty hours later, the low-molecular-weight DNA was isolated (18) and analyzed for AAV DNA replication by the Southern blotting procedure.

To determine AAV capsid antigen synthesis, the cells were harvested 40 h after transfection, resuspended in radioimmunoprecipitation assay buffer (50 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1% Triton X-100, 1% deoxycholate, 0.1% sodium dodecyl sulfate [SDS]) supplemented with 2 mM phenylmethylsulfonyl fluoride and 1 μg each of antipain, leupeptin, and pepstatin A, and centrifuged briefly at 4°C, and the supernatant fraction was kept at −20°C before analysis on an SDS–8% polyacrylamide gel electrophoresis gel. After electrophoresis, the polypeptides were transferred to a nitrocellulose membrane and immunodetection was performed with polyclonal capsid antibodies (kindly provided by Mertyn Malkinson) and a peroxidase label conjugate (Sigma). The chromogenic substrate 3-amino-9-ethylcarbazole (Sigma) was used to detect peroxidase activity.

To assay for infectious virus production, C17 cells were harvested 40 h post-transfection and lysed by several cycles of freeze-thawing. The extract was heated for 2 h at 56°C to inactivate adenovirus, clarified by centrifugation, mixed with a fresh sample of Ad2, and adsorbed to HeLa cells. At 40 h postinfection, the cells were lysed and Ad2 was inactivated as described above. One-tenth of the lysate was treated with DNase I (10 μg/ml) for 30 min at 37°C, denatured with 0.8 M NaOH, neutralized with 0.8 M $NH_4C_2H_3O_2$, dot blotted to nitrocellulose, and hybridized with a ^{32}P-labeled AAV DNA probe.

RESULTS

A useful feature of the EBV shuttle vector system is that the recombinant products can be retrieved in *E. coli* for structural analysis. Accordingly, we isolated EBV-AAV recombinants from cell lines that carried p220.2 shuttle vectors with 8.2-, 1.6-, and 0.51-kb segments of AAVS1 DNA and had been infected at various passage levels. The recombinants were analyzed with respect to AAV DNA content, the sequences at the junctions, and the capacity of the AAV genome to be rescued from the episomal shuttle vector by adenovirus infection.

AAV DNA in recombinants. To define the AAV sequences present on the EBV vectors after infection of the C17-p220.2 (AAVS1 kb 0-8.2), C17-p220.2 (AAVS1 kb 0-1.6), and C17-p220.2(AAVS1 kb 0-0.51) cell lines, we used a set of AAV oligonucleotides spanning the AAV genome to hybridize dot blots of the recombinant DNAs (Fig. 1). Some of the recombinants (9 of 43) reacted positively to all of the probes used for hybridization, suggesting that the entire AAV genome might be present. No evidence of the presence of several separate inserts of AAV DNA or several copies of the same DNA was

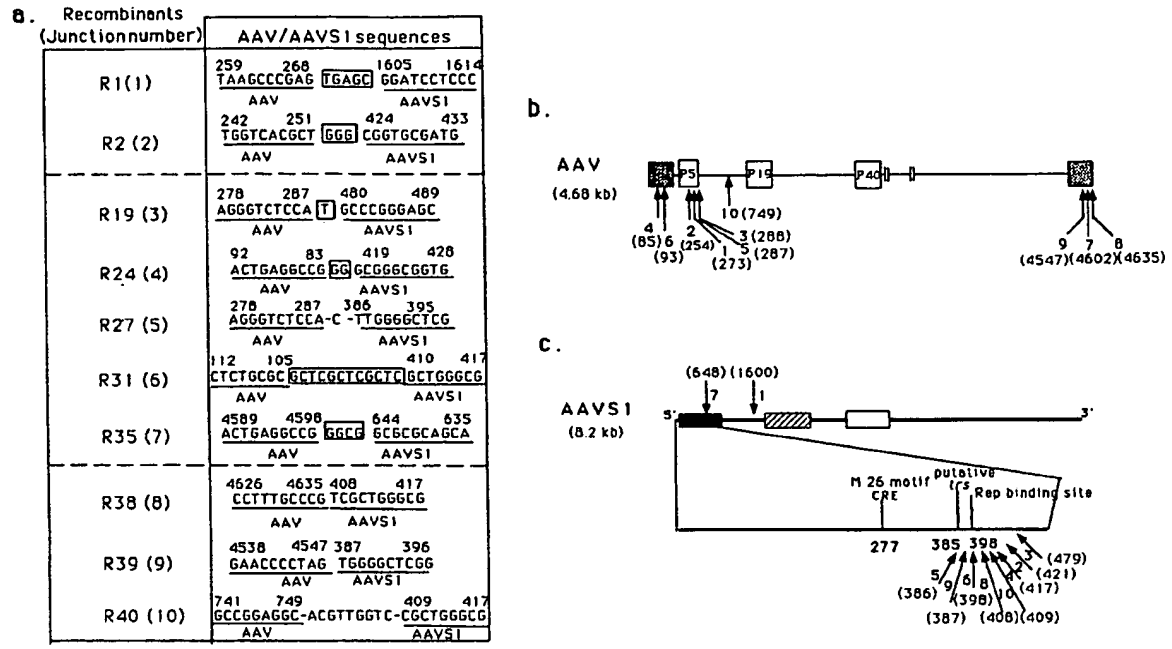

FIG. 2. Junctions between AAV and AAVS1 sequences. (a) Sequences at the nucleotide level. Boxed sequences at the junctions are common to both viral and vector genomes. A junction number (in parentheses) was given to each recombinant, and the same number is used in panels b and c. The horizontal dotted lines separate the three groups of recombinants isolated from three different cell lines (Fig. 1). (b) Schematic representation of the junctions with respect to the AAV genome. (c) Schematic representation of the junctions with respect to the AAVS1 segment of chromosome 19. The AAVS1 sequence (8.2 kb) shows a CpG island (▩), a region corresponding to a partial cDNA clone (▨), and a minisatellite repetitive DNA sequence (▭). The enlargement of the first 510 nt at the 5′ end of AAVS1 shows the M26 motif from *Saccharomyces cerevisiae*, a putative terminal resolution site (*trs*), and a Rep binding site. CRE, cyclic Amp response element.

found by restriction digestion (data not shown). Most of the recombinants (32 of 43) had deletions in the *rep* genes; the deletions extended to the capsid genes in at least 20 of these cases. Thirty-six of forty-three recombinants reacted positively to the oligonucleotide AAVD located in the D region of the itr. The D region of the AAV itr is outside of the 125-bp symmetrical sequences able to fold into a T-shaped structure. Hence, the itr, or at least part of it, is present in most of the recombinants. The pattern of the AAV sequences found in each recombinant (Fig. 1) was used to select AAV oligonucleotides for sequencing of the junctions between the viral and vector DNA segments.

AAV and vector sequences at recombinant junctions. The junctions of 11 recombinants (noted by asterisks in Fig. 1) were sequenced by using multiple oligonucleotides as primers (see Material and Methods and Fig. 2 to 5).

(i) Junctions between AAV and AAVS1 sequences. A junction between AAV DNA and the AAVS1 DNA segment of the p220.2 vector was found in all of the recombinants analyzed but one (R23; Fig. 2). Surprisingly, a second junction between these two DNAs was never found in the same recombinant. In AAV, the junctions were localized in the itr or near the P5 promoter (Fig. 2a and b). In one case (R40), the junction was found in the *rep* gene sequence.

In 80% of the cases, the junctions in the AAVS1 sequence were found in the 5′ 510-bp segment previously defined as the minimum chromosome 19 DNA sequence able to direct AAV integration (12). Strikingly, the junctions within the 510-bp segment were tightly clustered around the AAVS1 Rep binding site (Fig. 2a and c). Only two junctions were found 3′ to the 510-bp segment (in R1 and R35). Because the AAVS1 DNA at the junction was found to be oriented towards the 3′ end of AAVS1 (R35 is the sole exception; Fig. 2) and because a second junction with AAVS1 DNA was never found in the same recombinant DNA molecule, we wished to determine if AAVS1 sequences 5′ to the insertion might have been preferentially lost during integration. Therefore, we used three AAVS1 oligonucleotides (nt 60 to 79, 248 to 265, and 453 to 435) to screen by dot blot hybridization for the presence of AAVS1 DNA 5′ to the junction with AAV DNA. The 5′ AAVS1 DNA sequences were found in 7 of 11 cases (Table 1). Thus, although AAVS1 DNA 5′ to the AAV junction was present in 64% of the recombinant vectors, this DNA segment

TABLE 1. AAVS1 sequences from the 5′ 510-bp fragment present in EBV recombinant vectors[a]

Recombinant (junction point in AAVS1 sequence [nt])	Hybridization with AAVS1 oligonucleotide:		
	1 (nt 60–79)	2 (nt 248–265)	3 (nt 453–435)
R1 (1600)	–	–	–
R2 (421)	–	–	+
R19 (479)	+	+	+
R23 (?)	+	+	+
R24 (417)	+	+	+
R27 (386)	–	–	+
R31 (398)	–	–	+
R35 (648)	+	+	+
R38 (408)	+	+	+
R39 (387)	+	+	+
R40 (409)	+	+	+

[a] The presence of sequences from the 5′ end of the preintegration locus was assessed by dot blot hybridization with oligonucleotide probes.

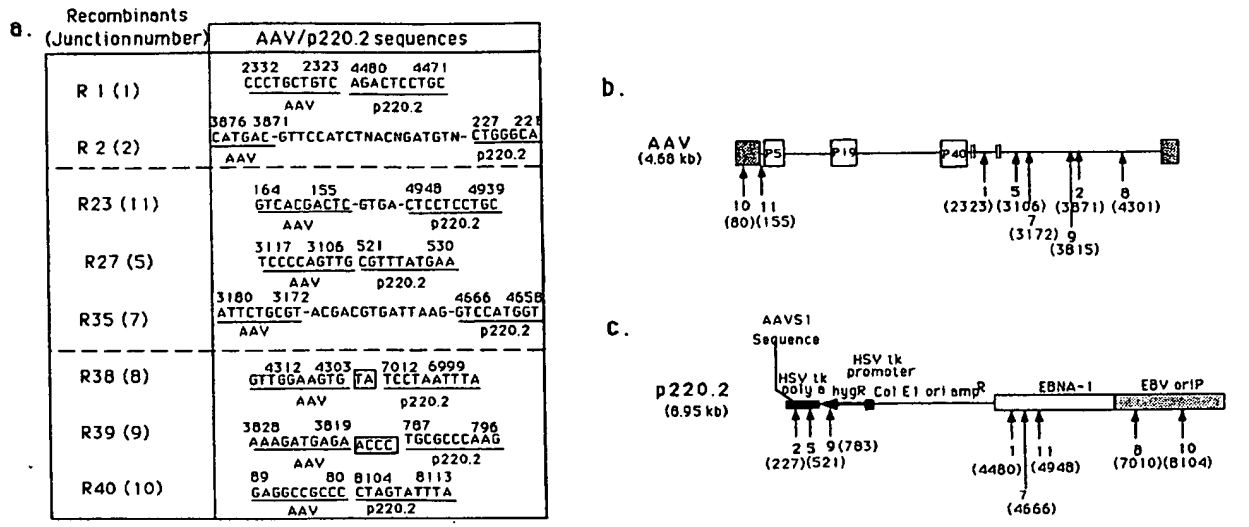

FIG. 3. Junctions between AAV and p220.2 sequences outside of the AAVS1 segment. (a) Sequences at the nucleotide level. Boxed sequences at the junctions are common to both viral and vector genomes. A junction number (in parentheses) was given to each recombinant, and the same number is used in panels b and c. The horizontal dotted lines separate the three groups of recombinants isolated from three different cell lines (Fig. 1). (b) Schematic representation of the junctions with respect to the AAV genome. (c) Schematic representation of the junctions with respect to the p220.2 plasmid genome. The 8.9-kb p220.2 plasmid contains the herpes simplex virus (HSV) thymidine kinase (tk) promoter and polyadenylation signal (poly a) flanking a hygromycin resistance-encoding gene (hygR), the EBV latent origin of replication (oriP), the EBV-encoded nuclear antigen EBNA-1, and *E. coli* sequences (Col E1 and ori amp^R).

was apparently translocated to other positions in the vector genome. Whether the rearrangements of AAVS1 DNA occurred during AAV integration or as a consequence of vector rearrangements which have been observed to occur prior to AAV infection (12) is not known, although the fact that the rearrangement always occurs immediately 5′ to the junction site suggests the former possibility.

Inspection of the AAV-AAVS1 junction sequences (Fig. 2a) revealed only limited patchy nucleotide sequence homology, indicating that integration into the vector occurs by nonhomologous recombination.

(ii) Junctions between AAV DNA and p220.2 vector DNA outside of the AAVS1 segment. As noted above, the second junction between AAV DNA and the vector genome occurred outside of the AAVS1 segment in each recombinant examined. With respect to the AAV genome, these junctions were located in the itr or in the *cap* gene (in the latter case, substantial portions of the *rep* and capsid genes were deleted in the recombinant genome) (Fig. 3a and b). With respect to the p220.2 vector, the junctions were found in the EBV *oriP* sequence, in the EBV EBNA-1 gene, and in the hygromycin resistance-encoding gene (Fig. 3a and c). Despite the crucial role of these vector elements for stable episome propagation in animal cells (27), the period of 48 h between AAV infection and recovery of the vectors in bacteria was probably too short for the cells to lose the plasmid by dilution. As expected, no junctions were found in the *E. coli* elements necessary for selection in bacteria by virtue of the screening procedure.

(iii) Head-to-tail junctions between AAV sequences. When primers located near the left and right itrs (AAV1 and AAV9; Fig. 1A) were used to sequence the junctions, it became evident that the itr and adjacent DNA were sometimes organized in a head-to-tail fashion, relative to the wild-type AAV genome (recombinants R1, R2, R19, R27, and R31 in Fig. 4). In the wild-type AAV genome, the first 125 nt of the 145-nt itr can form a T-shaped hairpin structure because of the presence of two small internal palindromes (B-B′ and C-C′) flanked by a larger palindrome (A-A′) (1). In addition, the complete itr element includes a single-stranded D region located outside of the palindromic hairpin. The head-to-tail AAV-AAV recombinant junctions are composed of one complete itr (D-A′-B′-B-C′-C-A) expanded by a 20-nt D′ region (D-A′-B′-B-C′-C-A-D′). Thus, in the recombinants with the head-to-tail orientation, the 3′ end of the AAV genome is joined to nt 126 at the 5′ end (Fig. 4). The R1, R2, R19, and R27 recombinants which display the unusual head-to-tail organization of the itr and adjacent DNA are also characterized by two additional common features; the junction with AAVS1 occurs near the P5 promoter, and a substantial segment of the AAV internal sequences (the *cap* and *rep* genes) is deleted (Fig. 2; see Fig. 6A). The fifth recombinant structure, which displays the head-to-tail organization of the itr and adjacent DNA at one end (R31), contains the internal AAV sequences and is linked to AAVS1 DNA via a disrupted itr at the other end (Fig. 6C); the AAV genome integrated into the R31 recombinant vector was rescued by adenovirus infection (see below).

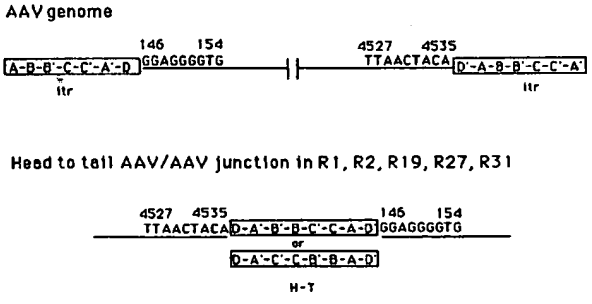

FIG. 4. Junctions between AAV sequences. The sequences D-A′-B′-B-C′-C-A-D′ and D-A′-C′-C-B′-B-A-D′ refer to the head-to-tail (H-T) organization of the recombinant AAV itr relative to the wild-type itr. The gap indicates a disruption in the sequence with respect to the scale of the drawing.

a.

Recombinants (Junction number)	AAV/unknown sequences
R19 (3)	1361 1352 TATGGCCTCC ?
R23 (11)	4612 4603 GACCTTTGGT ?
R24 (4)	4586 4595 CTCACTGAGG ?
R31 (6)	278 287 AGGGTCTCCA ?

b.

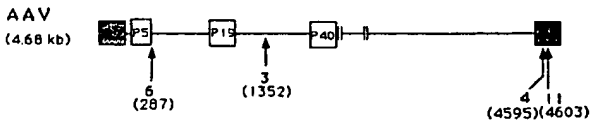

FIG. 5. Junctions between AAV and nonviral-nonvector sequences. (a) Sequences at the nucleotide level. A junction number (in parentheses) was given to each recombinant, and the same number is used in panel b. (b) Schematic representation of the junctions with respect to the AAV sequence.

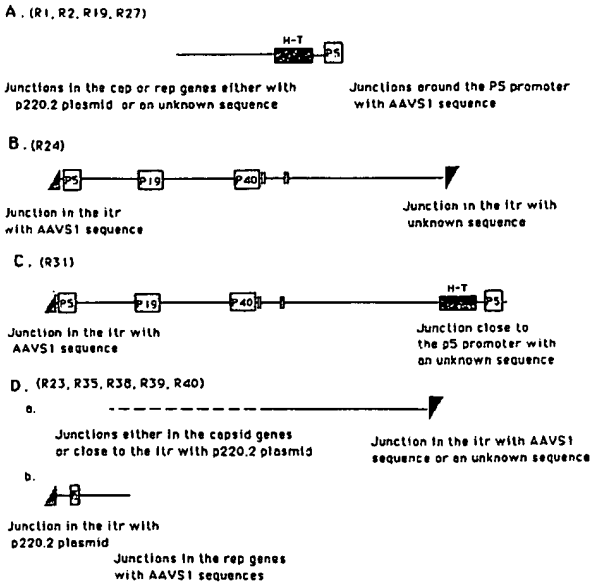

FIG. 6. Patterns of AAV inserts and junctions in EBV recombinant vectors. ▩, head-to-tail (H-T) organization of the itr (see the text for details) and adjacent DNA; ▼, disrupted itrs.

The head-to-tail organization of AAV DNA in recombinant R1, R2, R19, R27, and R31 DNAs was confirmed by PCRs with oligonucleotides AAV1 and AAV8 (Fig. 1A) as primers. The product was 230 nt long as judged by gel electrophoresis and hybridized with the AAV9 oligonucleotide (Fig. 1), thus confirming the presence of itr DNA.

When oligonucleotide primers AAV1 and AAV8 were used independently in PCRs, no products were detected, indicating that the organization of these AAV-AAV recombinant junctions is not head to head or tail to tail. The recombinants showing head-to-tail organization of the AAV itr and adjacent DNA were isolated in independent experiments from two different infected cell lines [R1 and R2 from C17-p220.2(AAVS1 kb 0-8.2) cells and R19 and R27 from C17-p220.2(AAVS1 kb 0-1.6) cells] (Fig. 1). The similarity of the head-to-tail organization of AAV DNA in each case is striking and suggests a common recombination-integration mechanism. The possibility that this involves circularization of AAV DNA prior to or during integration is discussed below.

(iv) Junctions between AAV DNA and DNA unrelated to vector or viral DNA. DNA unrelated to the vector or viral genomes was found at the junctions with AAV DNA in four recombinants (R19, R23, R24, and R31; Fig. 5). The unrelated sequence in R24 shows some homology with the DNA of the human Line-1 element (long interspersed repetitive sequence derived from a retrotransposon) (19). An oligonucleotide prepared from the R24 unrelated sequence hybridized with the DNAs of the R19 and R23 recombinants (but not with R31 DNA). All of the recombinants with nonviral-nonvector DNA, related to human Line-1 DNA, were derived from a single infected cell line propagating the p220.2(AAVS1 0-1.6 kb) episome. One vector isolate of this cell line which did not contain AAV DNA also hybridized with the oligonucleotide derived from the R24 nonviral-nonvector segment. These results suggest that some p220.2(AAVS1 0-1.6 kb) episomes acquired the human Line-1 DNA prior to recombination with AAV DNA. The human Line-1 element is known to be capable of existing as part of an extrachromosomal circular element (20).

Summary of recombinant junctions. On the basis of the sequencing data from 11 recombinants, the junctions can be grouped as shown in Fig. 6. In group A, one junction with the episome joins AAV DNA from around the P5 promoter to the episomal AAVS1 segment; the second junction occurs either between a p220.2 sequence (that is outside the AAVS1 segment) and the AAV capsid genes (R1, R2, and R27) or between an AAV *rep* gene sequence and nonviral-nonvector DNA (probably Line-1 DNA) (R19). All of the recombinants in group A show head-to-tail organization of the AAV itr and adjacent sequences, and much of the internal AAV sequences are deleted. In group B (R24), all of the internal AAV sequences appear to be present; one junction with AAVS1 is via a disrupted itr element; the second junction with human Line-1 DNA in the episome also occurs via a disrupted itr element. In group C (R31), the organization of the AAV insert is similar to that of group B except that one junction shows the head-to-tail organization of the viral itr and adjacent DNA and connects with episomal nonviral-nonvector DNA at a point near the P5 promoter. The AAV insert in R31 was rescuable by adenovirus infection (see below). Group D (R23, R35, R38, R39, and R40) is characterized by the presence of a segment of AAV DNA and a disrupted itr joined to either AAVS1, p220.2, or the nonviral-nonvector DNA.

Rescue of AAV from EBV recombinant vectors. It was of interest to determine if the AAV DNA in the recombinants which reacted with all of the virus-specific oligonucleotide probes (R4, R5, R9, R17, R23, R24, R26, R29, and R31 in Fig. 1) could be rescued by infection with adenovirus. Accordingly, C17 cells were first infected with Ad2 and then transfected with the plasmid-purified shuttle vector-AAV recombinants noted above. At 40 h posttransfection-infection, expression of the AAV capsid genes, rescue and replication of the AAV DNA, and production of infectious particles from these EBV recombinant vectors were assessed.

AAV capsid gene expression was determined by Western blot (immunoblot) analysis with a polyclonal AAV capsid antibody (Fig. 7). R31 and R26 (data not shown) expressed the capsid proteins at a level similar to that seen with the pSM620 plasmid (34), suggesting that the AAV DNA is rescued from the EBV recombinant vectors and replicated. A low expression

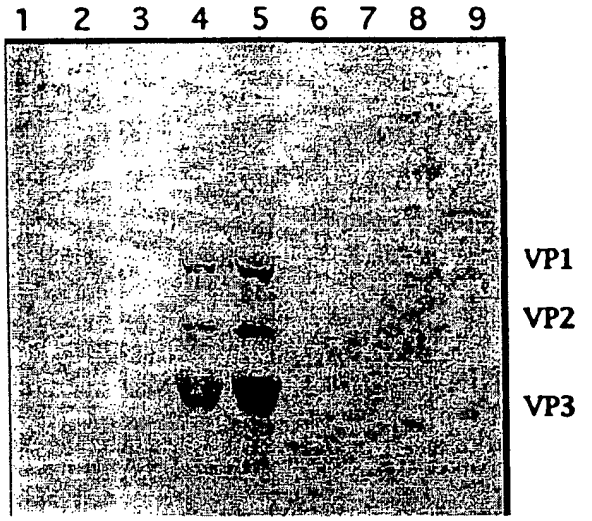

FIG. 7. Capsid gene expression of AAV rescued from EBV recombinant vectors. Immunoblot analysis of C17 cells infected with Ad2 and transfected with plasmid pAAV/Ad (35) (lane 3), plasmid pSM620 (34) (lane 4), recombinant R31 (lane 5), recombinant R9 (lane 6), recombinant R24 (lane 7), or recombinant R23 (lane 8). Lanes 1 and 9 contained, respectively, a Rainbow molecular weight marker (Amersham) and a prestained SDS-polyacrylamide gel electrophoresis standard high-range marker (Bio-Rad). The molecular masses of proteins VP1, VP2, and VP3 are 87, 73, and 62 kDa. Lane 2 contained a mock-transfected control. The antiserum used in the immunoblots was a polyclonal anti-AAV capsid serum.

level of the VP polypeptides was seen with some of the other recombinants (Fig. 7).

To assess rescue and DNA replication, the low-molecular-weight DNA was extracted from the transfected-infected C17 cells and analyzed on a 0.7% agarose gel (Fig. 8A). Only the extracts from cells transfected with recombinants R31 and R26 (data not shown) contained detectable amounts of the AAV monomer and dimer forms also seen when AAV DNA is rescued from the pSM620 plasmid. These products were resistant to *Dpn*I digestion (data not shown). Hybridization with an AAV probe (Fig. 8B) confirmed the nature of the DNA.

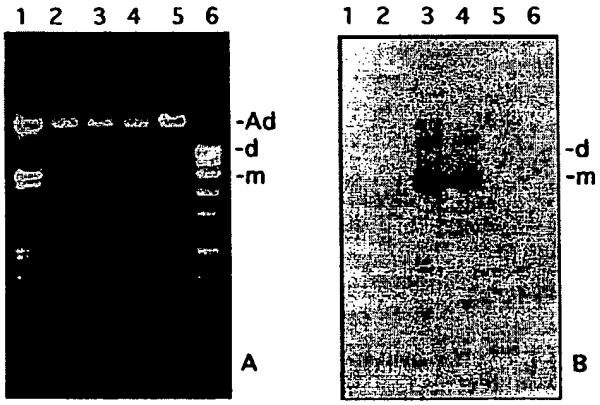

FIG. 8. Rescue and replication of AAV DNA from EBV recombinant vectors. Extrachromosomal DNA was extracted from C17 cells infected with Ad2 and transfected with recombinant R9 (lane 2), recombinant R31 (lane 3), or plasmid pSM620 (lane 4) or mock transfected (lane 5). Lanes 1 and 6 contained DNA molecular weight markers III and VII (Boehringer Mannheim) (A) Agarose (0.7%) gel stained with ethidium bromide. (B) Southern blot hybridization with an AAV probe. Abbreviations: m, monomer, d, dimer; Ad, adenovirus.

Production of infectious particles was assessed by adsorbing the lysates of transfected-infected C17 cells to adenovirus-infected HeLa cells. At 40 h postinfection, the presence of progeny AAV in the HeLa cells was determined by dot blot hybridization (AAV probe) of HeLa cell lysates treated first with DNase I (to screen for DNase-resistant particles) and then with NaOH to extract and denature the DNA. Two recombinants (R26 and R31) gave rise to progeny AAV in the above-described procedure. Both R26 and R31 contain the head-to-tail organization of the itr and adjacent sequences. Thus, despite this divergence from the wild-type organization of the AAV genome, the AAV inserts in R26 and R31 could be rescued and replicated to infectious particles.

DISCUSSION

AAV is unique among animal viruses in its capacity to undergo site-specific chromosomal integration. To aid the investigation of the targetting mechanism, we previously developed an EBV-based shuttle vector system in which the chromosome 19 AAVS1 preintegration DNA propagates as an episome into which AAV can integrate (12). The objective of the present experiments was to define the structure of the recombinant junctions in episomal integration and to compare such junctions with AAV insertions into AAVS1 when it is part of the intact chromosome.

Although head-to-tail and tail-to-tail arrangements in chromosomal inserts containing multiple copies of the AAV genome have been reported previously (5, 22, 26, 30, 45), the finding of head-to-tail organization of the AAV itr and adjacent sequences in 5 of 11 recombinants sequenced (Fig. 6A and C) was not anticipated because only 1 of these contained an insert of greater than unit length. In these recombinants, the itr was modified by addition of a 20-nt D region and linked the 3′ end of the AAV genome to AAV nt 126 at the 5′ end. In all five of these recombinant junctions, the crossover point with the vector occurred at or slightly downstream of the P5 promoter at AAV nt 250 (Fig. 2b). Although the retention of viral DNA in this group of independently isolated recombinants varied from a rescuable, complete genome equivalent in R31 to extensive deletions in the *rep* and *cap* genes in R1, R2, R19, and R27, the structure of the head-to-tail modification was remarkably similar in each case. Multiple copies of the AAV genome in tandem array could occur in an insert as the consequence of either recombination or replication of the inserted sequence before or during integration. Recombination via the itr could result in either head-to-tail, tail-to-tail, or head-to-head orientation of adjacent copies of the genome. The former two possibilities have been reported (5, 16, 22, 26, 30, 45). Although the predominant frequency of head-to-tail orientation in chromosomal integration suggests the likelihood that a rolling-circle form of replication (11) is involved; even in this event, the first step would be formation of a circular intermediate by recombination via the itr. Integration of less than a full genome equivalent of any of the recombinants characterized in this report does not require replication but does suggest that a circular molecule is an intermediate in the process. As discussed elsewhere, (31), if such integrated forms do arise by replication, the mode of replication must differ from that of AAV during lytic infection, which gives rise only to head-to-head or tail-to-tail intermediates (1, 17, 40). The proposal that viral genomes may undergo an alternative mode of replication, quite different from that characteristic of replication during lytic infection, is not without precedent. Aberrant rolling-circle

replication intermediates—with high recombinogenic potential—have been described for papovavirus DNA (2, 7, 15), which normally replicates bidirectionally, giving rise to theta-type intermediates during lytic infection (41). Indeed, rolling-circle replication intermediates have been proposed as intermediates in papovavirus integration (8).

Previously we have shown that a 510-bp segment at the 5′ end of AAVS1 that has been cloned (23) is capable of directing site-specific integration of AAV into the EBV-based episomal vector (12). Likely recognition signals that might target the AAV genome to AAVS1 are the Rep binding site at AAVS1 nt 398 to 413 (6, 42), a potential terminal resolution site (38) at nt 384 to 389, and the yeast recombinogenic M26 motif (32, 37) at nt 277 to 283. Our analysis of the recombinant junctions in the episome has shown that in 80% of the cases examined, the AAV insertion points are tightly clustered around the Rep binding site (AAVS1 nt 386 to 479; Fig. 2c). However, two junctions with AAVS1 DNA, both 3′ to the 510-bp segment (AAVS1 nt 648 and 1600), were also detected. The tight clustering of a major proportion of crossover points around the potential terminal resolution site and Rep binding motif certainly underlines the importance of these signals in the targetting mechanism.

In latently infected cell lines, the few AAV insertion points that have been determined were located at AAVS1 nt 1026 to 1030 and 1144 to 1146 (23), in a region from nt 713 to 1303 (36, 45), and at position 727 (13). These insertion points in AAVS1, when it is part of intact chromosome 19, are somewhat 3′ to the tight cluster around positions 386 to 479 when AAVS1 is propagated in an episomal vector. The reason for this difference is not known. It should be noted, however, that mapping of AAVS1 recombinant junctions in latently infected cells can be performed only after many cell generations of growth subsequent to the initial integration event. In contrast, the recombinants generated in the EBV episomal system are not subject to this constraint.

Apart from the difference in the distance of the insertion points from the Rep binding motif, noted above, integration of AAV into AAVS1 in the manipulatable EBV-based episomal system shows many points of similarity to AAV integration into AAVS1 as part of intact chromosome 19. A genome equivalent of the viral DNA can be rescued by adenovirus superinfection in both cases. The AAVS1 target is substantially disrupted and rearranged both in the episome and in the chromosome. In the episome, AAVS1 sequences 5′ to the viral insertion are no longer present at the junction, having apparently been translocated to other vector positions. In the chromosome, disruption of AAVS1 is severe enough that it becomes diagnostic for AAV site-specific integration (25). The head-to-tail organization of the viral itr and adjacent DNA can also be encountered in both chromosomal integration (45) and episomal integration. Similarly, viral recombinant junctions localized to the P5 promoter region (which also contains a Rep binding sequence; 28, 29) have been noted in both systems (45). Thus, the highly manipulatable EBV-based episomal system reflects many aspects of AAV integration at the intact-chromosomal level.

ACKNOWLEDGMENTS

We thank Erik Falck-Pedersen, Peter Ward, Saburo Kashii, and Michael Linden for helpful discussions and critical reading of the manuscript. We thank N. Cortez for excellent technical assistance.

This work was supported by grant AI122251 from the U.S. Public Health Service.

REFERENCES

1. Berns, K. I. 1990. Parvovirus replication. Microbiol. Rev. 54:316–329.
2. Bjursell, G. 1978. Effects of 2′-deoxy-2′-azidocytidine on polyoma virus DNA replication: evidence for rolling circle-type mechanism. J. Virol. 26: 136–142.
3. Canfield, V., J. R. Emanuel, N. Spickofsky, R. Levenson, and R. F. Margolskee. 1990. Ouabain-resistant mutants of the rat Na,K-ATPase α2 isoform identified by using an episomal expression vector. Mol. Cell. Biol. 10:1367–1372.
4. Carter, B. J. 1992. Adeno-associated virus vectors. Curr. Opin. Biotechnol. 3:533–539.
5. Cheung, A. K., M. D. Hoggan, W. W. Hauswirth, and K. I. Berns. 1980. Integration of the adeno-associated virus genome into cellular DNA in latently infected human Detroit 6 cells. J. Virol. 33:739–748.
6. Chiorini, J. A., M. D. Weitzman, R. A. Owens, E. Urcelay, B. Safer, and R. M. Kotin. 1994. Biologically active Rep proteins of adeno-associated virus type 2 produced as fusion proteins in *Escherichia coli*. J. Virol. 68:797–804.
7. Deichaite, I., Z. Laver-Rudich, D. Dorsett, and E. Winocour. 1985. Linear simian virus 40 DNA fragments exhibit a propensity for rolling-circle replication. Mol. Cell. Biol. 5:1787–1790.
8. Dorsett, D., I. Deichaite, and E. Winocour. 1985. Cellular and linear simian virus 40 DNAs differ in recombination. Mol. Cell. Biol. 5:869–880.
9. DuBridge, R. B., P. Tang, H. C. Hsia, P.-M. Leong, J. H. Miller, and M. P. Calos. 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol. Cell. Biol. 7:379–387.
10. Felgner, P. L., T. R. Gadek, M. Holm, R. Roman, H. W. Chan, M. Wenz, J. P. Northrop, G. M. Ringold, and M. Danielsen. 1987. Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. Proc. Natl. Acad. Sci. USA 84:7413–7417.
11. Gilbert, W., and D. Dressler. 1968. DNA replication: the rolling circle. Cold Spring Harbor Symp. Quant. Biol. 33:473–484.
12. Giraud, C., E. Winocour, and K. I. Berns. 1994. Site-specific integration by adeno-associated virus is directed by a cellular DNA sequence. Proc. Natl. Acad. Sci. USA 91:10039–10043.
13. Goodman, S., X. Xiao, R. E. Donahue, A. Moulton, J. Miller, C. Walsh, N. S. Young, R. J. Samulski, and A. W. Nienhuis. 1994. Recombinant adeno-associated virus-mediated gene transfer into hematopoietic progenitor cells. Blood 84:1492–1500.
14. Greener, A. 1990. E. coli SURE TM strain: clone "unclonable" DNA. Strategies 3:5–6.
15. Grossman, Z., K. I. Berns, and I. Winocour. 1985. Structure of simian virus 40-adeno-associated virus recombinant genomes. J. Virol. 56:457–465.
16. Handa, H., K. Shiroki, and H. Shimojo. 1977. Establishment and characterization of KB cells lines latently infected with adeno-associated virus type 1. Virology 82:84–92.
17. Hauswirth, W. W., and K. I. Berns. 1979. Adeno-associated virus DNA replication: non unit-length molecules. Virology 93:57–68.
18. Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. 26:365–369.
19. Hohjoh, H., R. Minakami, and Y. Sakaki. 1990. Selective cloning and sequence analysis of the human L1 (Line-1) sequences which transposed in the relatively recent past. Nucleic Acids Res. 18:4099–4104.
20. Jones, R. S., and S. Potter. 1985. L1 sequences in Hela extrachromosomal circular DNA: evidence for circularization by homologous recombination. Proc. Natl. Acad. Sci. USA 82:1989–1993.
21. Kotin, R. M. 1994. Prospects for the use of adeno-associated virus as a vector for human gene therapy. Hum. Gene Ther. 5:793–801.
22. Kotin, R. M., and K. I. Berns. 1989. Organization of adeno-associated virus DNA in latently infected Detroit 6 cells. Virology 170:460–467.
23. Kotin, R. M., R. M. Linden, and K. I. Berns. 1992. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. EMBO J. 11:5071–5078.
24. Kotin, R. M., J. C. Menninger, D. C. Ward, and K. I. Berns. 1991. Mapping and direct visualization of a region-specific viral DNA integration site on chromosome 19q13-qter. Genomics 10:831–834.
25. Kotin, R. M., M. Siniscalco, R. J. Samulski, X. D. Zhu, L. Hunter, C. A. Laughlin, S. McLaughlin, N. Muzyczka, M. Rocchi, and K. I. Berns. 1990. Site-specific integration by adeno-associated virus. Proc. Natl. Acad. Sci. USA 87:2211–2215.
26. Laughlin, C. A., C. B. Cardellichio, and H. C. Coon. 1986. Latent infection of KB cells with adeno-associated virus type 2. J. Virol. 60:515–524.
27. Margolskee, R. F. 1992. Epstein-Barr virus based expression vectors. Curr. Top. Microbiol. Immunol. 158:67–95.
28. McCarty, D. M., D. J. Pereira, I. Zolotukhin, X. Zhou, J. H. Ryan, and N. Muzyczka. 1994. Identification of linear DNA sequences that specifically bind the adeno-associated virus Rep protein. J. Virol. 68:4988–4997.
29. McCarty, D. M., J. H. Ryan, S. Zolotukhin, X. Zhou, and N. Muzyczka. 1994. Interaction of the adeno-associated virus Rep protein with a sequence within the A palindrome of the viral terminal repeat. J. Virol. 68:4998–5006.
30. McLaughlin, S. K., P. Collis, P. L. Hermonat, and N. Muzyczka. 1988. Adeno-associated virus general transduction vectors: analysis of proviral structures. J. Virol. 62:1963–1973.
31. Muzyczka, N. 1992. Use of adeno-associated virus as a general transduction vector for mammalian cells. Curr. Top. Microbiol. Immunol. 158:97–129.
32. Ponticelli, S. A., and G. R. Smith. 1992. Chromosomal context dependence

of a eukariotic recombinational hot spot. Proc. Natl. Acad. Sci. USA 89:227–231.
33. Samulski, R. J. 1993. Adeno-associated virus: integration at a specific chromosomal locus. Curr. Opin. Genet. Dev. 3:74–80.
34. Samulski, R. J., K. I. Berns, M. Tan, and N. Muzyczka. 1982. Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. Proc. Natl. Acad. Sci. USA 79:2077–2081.
35. Samulski, R. J., L. S. Chang, and T. Shenk. 1989. Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. J. Virol. 63:3822–3828.
36. Samulski, R. J., X. Zhu, X. Xiao, J. D. Brook, D. E. Housman, N. Epstein, and L. A. Hunter. 1991. Targeted integration of adeno-associated virus (AAV) into human chromosome 19. EMBO J. 10:3941–3950.
37. Schuchert, P., M. Langsford, E. Kaslin, and J. Kohli. 1991. A specific DNA sequence is required for high frequency of recombination in the ade 6 gene of fission yeast. EMBO J. 10:2157–2163.
38. Snyder, R. O., D.-S. Im, T. Ni, X. Xiao, R. J. Samulski, and N. Muzyczka. 1993. Features of the adeno-associated virus origin involved in substrate recognition by the viral Rep protein. J. Virol. 67:6096–6104.
39. Stamatatos, L., R. Leventis, M. J. Zuckermann, and J. R. Silvius. 1988. Interactions of cationic lipid vesicles with negatively charged phospholipid vesicles and biological membranes. Biochemistry 27:3917–3925.
40. Straus, S. E., E. D. Sebring, and J. A. Rose. 1976. Concatemers of alternating plus and minus strands are intermediates in adenovirus-associated virus DNA synthesis. Proc. Natl. Acad. Sci. USA 73:742–746.
41. Tooze, J. (ed.). 1980. Molecular biology of tumor viruses. 2nd ed. Part 2. DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
42. Weitzman, M. D., S. R. Kyostio, R. M. Kotin, and R. A. Owens. 1994. Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. Proc. Natl. Acad. Sci. USA 91:5808–5812.
43. Xiao, X., W. deVlaminck, and J. Monahan. 1993. Adeno-associated virus (AAV) vectors for gene transfer. Adv. Drug Delivery Rev. 12:201–215.
44. Yates, J. L., N. Warren, and B. Sugden. 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. Nature (London) 313:812–815.
45. Zhu, X. 1993. Characterization of adeno-associated virus proviral structure in latently infected human cells. Ph.D. thesis. University of Pittsburgh, Pittsburgh, Pa.